Design, Synthesis and Characterization of "*Janus*-PNAs": Novel Backbone Analogues That Form Double Duplexes With cDNA

A Thesis

Submitted in partial fulfillment of the requirements

of the degree of

Doctor of Philosophy

By

Manoj Kumar Gupta

ID: 20143336

IISER PUNE

INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH, PUNE

July 2019

This Thesis dedicated to….

My Supervisor

And

My Mother

Prof. Krishna N. Ganesh
FNA, FASc, FTWAS
Director. IISER Tirupati
Co-Editor: ACS Omega
Honorary Faculty, IISER Pune

CERTIFICATE

Certified that the work incorporated in the thesis entitled **Design, Synthesis and Characterization of "*Janus*-PNA": Novel Backbone Analogues That Form Double Duplexes with cDNA** submitted by **Mr. Manoj Kumar Gupta** was carried out by the candidate, under my supervision. The work presented here or any part of it has not been included in any other thesis submitted previously for the award of any degree or diploma from any other university or institution.

Date: 31ˢᵗ July 2019, Pune

Prof. Krishna N. Ganesh
(Research Supervisor)

DECLARATION

I declare that, this written submission represents my ideas in my own words and where other's ideas have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea / data / fact / source in my submission. I understand that violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Date: 31st July 2019, Pune   **Mr. Manoj Kumar Gupta**

                                                                                                ID: 20143336

Acknowledgements

Table of Contents

Chapter 4: Biophysical Studies of *Janus* PNA Oligomers

Chapter 5: i-Motif and G-Quadruplex Structures from *Janus* PNA Oligomers and their Morphological study by FESEM

Abbreviations

A	Adenine
Abs.	Absolute
AcOH	Acetic acid (glacial)
Ac$_2$O	Acetic anhydride
ACN	Acetonitrile
aeg	Aminoethylglycine
eaz	Ethylazide
tz	Triazolyl
etz	Ethyltriazolyl
ap	Antiparallel
aq.	Aqueous
(Boc)$_2$O	Boc anhydride
Bn	Benzyl
BIAB	(Diacetoxyiodo)benzene
Bt	Benzotriazole
C	Cytosine
Calcd	Calculated
Cbz	Benzyloxycarbonyl
CD	Circular Dichroism
CHCA	α-cyano-4-hydroxycinnamic acid
CuAAC	Copper mediated azide-alkyne cycloaddition
DCC	Dicyclohexylcarbodiimide
DCM	Dichloromethane
DHB	2,5-dihydroxybenzoic acid
DHBtOH	3-Hydroxy-1,2,3-benzotriazin-4(3H)-one
DIPEA	N,N-Diisopropylethylamine
DMAP	N,N-Dimethyl-4-aminopyridine
DMF	N,N-dimethylformamide
DMSO	N,N-Dimethyl sulfoxide

DNA	2'-deoxyribonucleic acid
ds	Double stranded
EBA	Ethylbromo acetate
EDTA	Ethylene diamine tetraacetic acid
Et	Ethyl
EtOAc	Ethyl acetate
FESEM	Field Emission Scanning Electron Microscope
g	gram
G	Guanine
gly	Glycine
h	Hours
HBTU	2-(1H-Benzotriazole-1-yl)- 1,1,3,3 tetramethyl-uronum-hexafluoro-phosphate
HOBt	N-Hydroxybenzotriazole
HPLC	High Performance Liquid Chromatography
HRMS	High Resolution Mass Spectrometry
in situ	In the reaction mixture
in vivo	Within the living
ITC	Isothermal Titration Calorimetry
IR	Infra-red
L-	Levo-
Lys	Lysine
MALDI-TOF	Matrix Assisted Laser Desorption Ionisation-Time of Flight
MBHA	4-Methyl benzhydryl amine
mg	milligram
MHz	Megahertz
μL	Microliter
μM	Micromolar
min	minutes
mL	milliliter

mM	millimolar
mmol	millimoles
mp	melting point
MW	Molecular weight
N	Normal
nm	Nanometer
NMR	Nuclear Magnetic Resonance
Obsvd	Observed
ONs	Oligonucleotides
p	Parallel
PCR	Polymerase chain reaction
Pd	Palladium
PIDA	iodosobenzene diacetate (phenyliodonium diacetate)
ppm	Parts per million
PNA	Peptide Nucleic Acid
PS-oligo	Phosphorothioate-oligo
R_f	Retention factor
RNA	Ribonucleic Acid
RP	Reverse Phase (-HPLC)
rt	Room temperature
RT	Retention time
S	Sinister
SPPS	Solid Phase Peptide Synthesis
ss	Single strand/single stranded
T	Thymine
TBAc	tert-Butyl bromoacetate
TEA/Et$_3$N	Et$_3$N/Triethylamine
TEMPO	Tetramethylpiperidine 1-oxyl
TFA	Trifluoroacetic acid
Tf$_2$O	Triflicanhydride

TfN$_3$	Trifluoromethanesulfonyl azide
TFMSA	Trifluoromethane sulfonic acid
THF	Tetrahydrofuran
TLC	Thin layer chromatography
T_m	Melting temperature
UV-Vis	Ultraviolet-Visible

Abstract

Design, Synthesis and Characterization of "*Janus* PNAs": Novel PNA backbone Analogues that form Double Duplexes with cDNA

The thesis entitled ''***Design, Synthesis and characterization of "Janus PNAs": Novel PNA backbone analogues that form double duplexes with cDNA***'' is comprised of studies towards the design, synthesis, biophysical evaluation of a new motif of peptide nucleic acid (PNA) analogs, *Janus* PNA. The modifications consist of incorporation of chiral (azide/triazole) side chains at α-position in the glycine segment of aminoethylglycine PNA backbone. These side chains carry a second nuclobase at each monomer leading to backbone *Janus* PNAs. The PNA oligomers synthesized on solid phase were investigated for their binding to target DNA by various biophysical techniques. The work also describes the i-Motif, G-quadruplex formation and stability of *Janus* PNAs and its hybridization studies with complementary DNA oligomers. The thesis is presented as five chapters.

Chapter 1: Introduction to PNA

Oligonucleotides (ONs) capable of sequence-specific recognition of nucleic acids (DNA/RNA) in specific and stable manner, are becoming more important for research, diagnosis and therapy. Binding of oligodeoxynucleotides (ODNs) to single stranded RNA or duplex DNA through triple-helix formation provides a way to modulate gene expression. The specific inhibition is based on the Watson-Crick base pairing between the heterocyclic bases of the antisense ONs and the target nucleic acid. For a successful medicinal agent, it is necessary for natural ONs to be chemically modified in a suitable manner. The chemical modifications of ONs have resulted in the synthesis and analysis of a large variety of oligonucleotide derivatives with modifications to the phosphate, the ribose or the nucleobase. This chapter gives an overview on the background literature on structure of DNA, and their analogues emphasizing the recent advancements in the field of peptide nucleic acids.

Peptide Nucleic Acids (PNA, Figure 1) introduced by Nielsen et al. are novel DNA mimics. It binds to natural oligonucleotides (DNA/RNA) in a sequence specific manner through complementary base pairing, resistant to both nucleases and proteases and has no known toxixity. Hence it has great potential for the development as antisense/antigene agents. The major

drawbacks of PNA like poor water solubility, inefficient cellular uptake, self-aggregation and ambiguity in directional selectivity of binding restrict its applications in terms of designing PNA-based gene-targeted drugs and showing weak supramolecular self-assembly. Hence, various modifications of PNA to overcome these limitations have been employed by different research groups. It also presents a survey of the literature relevant to the area of nucleic acid therapeutics. The effect of the different structural modifications on biophysical and biochemical properties of PNA and bifacial nucleobase is overviewed to draw directions for the *Janus* PNA work, aiming towards exploring new pathways of PNA design and applications.
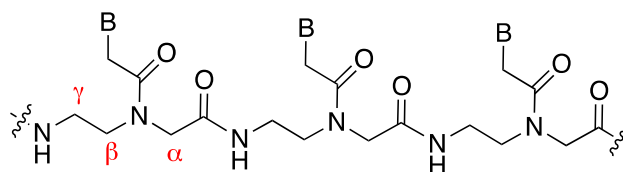
Peptide Nucleic acid

Present work: The work presented in this thesis deals with a new and novel design of dual binding two strands of natural oligonucleotides by a single strand of PNA and these are termed as *Janus* PNA (Figure 2). The synthesis, characterization and biophysical studies of such *Janus* PNA where nucleobases are anchored on two sides of a single PNA backbone are presented in this thesis. The hybridization properties of such PNAs to form duplexes, double duplexes and triplexes are demonstrated. The conformational features of such complexes examined by CD and the thermodynamic features of their hybrids studied by ITC are reported. Further exploration of i-motif and G-tetraplexes from such *Janus* PNAs and morphological features of their supramolecular assembly are studied.

Janus PNA

x
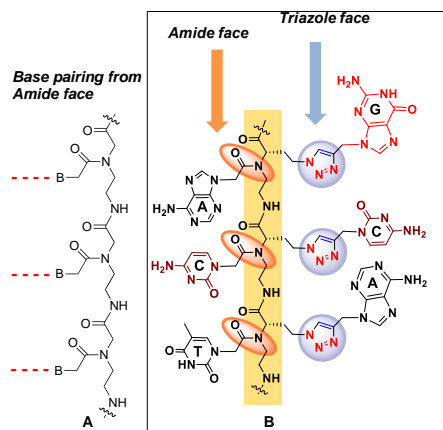
Chapter 2: Synthesis and characterization of C$_\alpha$ substituted *aeg*PNA monomer precursors for assembly of designed *Janus* PNAs

This chapter describes synthesis of precursor monomers required for assembling various *Janus* PNA oligomers. These are based on C$_\alpha$-ethylazido substitution that are useful for by post-assembly click reaction to generate desired *Janus* homo-oligomers. Such C$_\alpha$-(*S*-ethylazido) aminoethylglycyl A/T/C/G PNA monomers (**1**) can be coupled with propargyl nucleobases in solution to obtain *Janus* monomers (eg. **2**) that can be directly used to assemble mixed sequence *Janus* PNA oligomers. Alternatively, they can be incorporated on solid phase and coupled with nucleobases through global click reaction using propargyl nucleobases. The *Janus* PNA monomers (**2**) containing combinations of preinstalled nucleobases on both sides were also used for assembling *Janus* PNA oligomers.
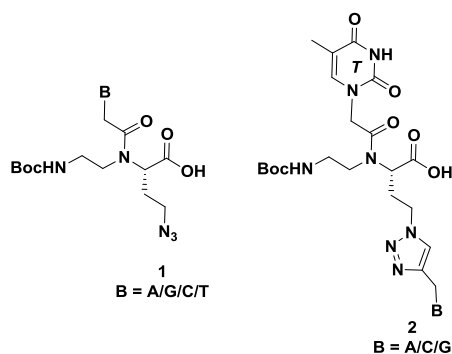
Chapter 3: Solid phase synthesis, purification and characterization of *Janus* PNA oligomers

The different *Janus* PNA monomers were used to assemble *Janus*-PNA oligomers by solid phase synthesis. The wavy line in structures represent PNA backbone.

Scheme 1 Synthetic strategy of *chimeric Janus* PNA on solid support
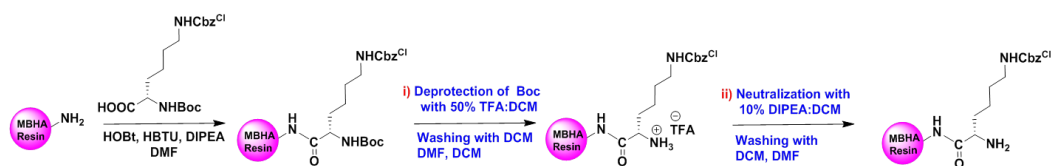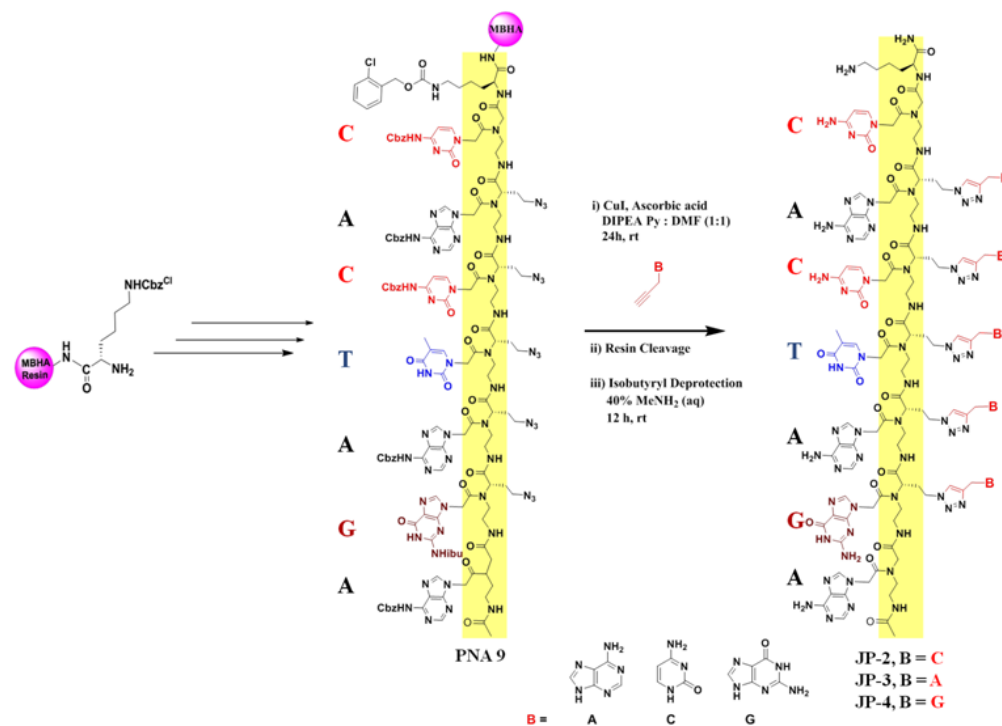
Scheme 2 Synthetic strategy of *Hetero Janus* PNA oligomer

P = MBHA resin
(a) Coupling reaction
(b) Click reaction
(c) Cleavage from resin
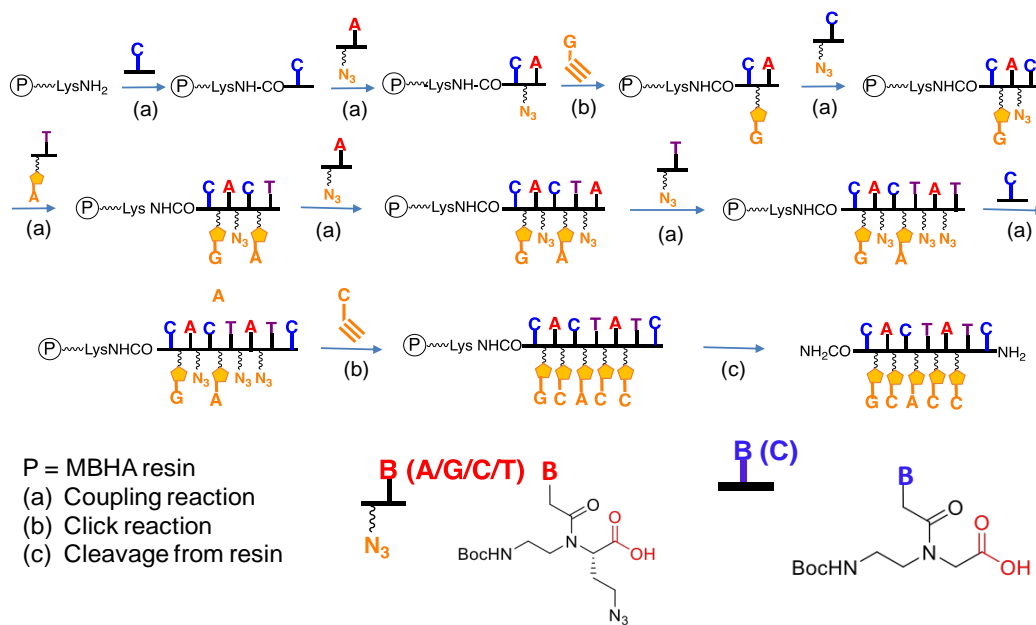
Various Kinds of *Janus* PNA oligomers
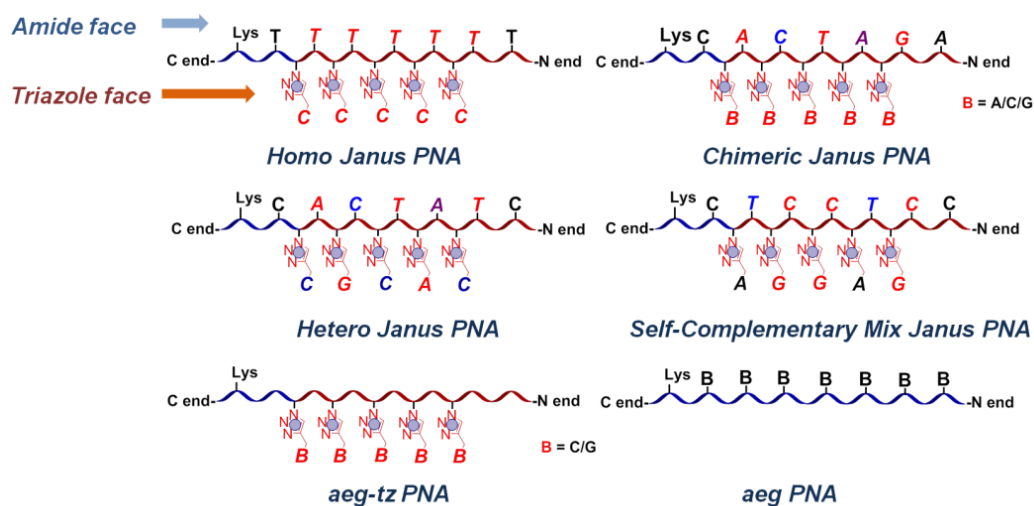
PNA analogues with nucleobase only on aminoethyl glycyl triazole face were synthesized to study hybridization exclusively from triazole and amide faces.

The different *Janus* PNAs, *aeg-tz* PNA, and *aeg* PNA oligomers synthesized as above by solid phase synthesis were cleaved from resin and purified by reverse phase HPLC followed by mass characterization using MALDI-TOF spectrometry..

Chapter 4: Biophysical Studies of *Janus* PNA Oligomers

This chapter describes the thermal stability of duplex, triplex and synergistic effect in double duplex or duplex of triplex by UV-T_m, sequential formation of complex by CD and thermodynamic study of various *Janus* PNA:DNA complexes with appropriate cDNA (Figure 11 to 15). The UV-T_m studies showed that the PNAs with nucleobases only on amide face or on triazolyl face show single transition as expected. *Janus* PNAs having nucleobases on both faces, form duplexes with appropriate complementary DNA on both faces leading to double duplexes or duplex of triplex, depending on the nature of sequences. Most interestingly, the T_m values of each duplex (amide face / triazole face) of double duplex from *Janus* PNAs are enhanced significantly over that of isolated duplexes individual faces with ΔT_m ranging from +15 ^{o}C to as much as +25 ^{o}C. Further, the T_m of the triazole face duplex in *Janus* PNAs with complementary DNA was always more than T_m of the amide face duplex ($\Delta T_m \sim$ +18 ^{o}C to +48 ^{o}C). This was a very surprising and unexpected but highly interesting and significant outcome of this work. The

amide face and triazole face duplexes also show characteristic CD patterns. Study of sequential formation of duplex from CD by changing the order of complementation, indicated that the order of formation of duplex does not matter and the conformation of the resulting double duplex is always same. Thus the most notable feature of double duplex formation by *Janus* PNAs is the synergistic enhancement in stability observed for each face duplex from the formation of duplex from the other face.

JP 1:DNA1:JP 1

PNA₂:DNA homoTriplex

JP 1:DNA 2

Triazole face homo duplex

DNA 2:JP 1:DNA 1:JP 1:DNA 2

Double triazole face duplex of homo amide face triplex

DNA 3:JP 2

Amide face heteroduplex

JP 2:DNA 2

Triazole face homo duplex

DNA 3:JP 2:DNA 2

hetero Amide, homo Triazole Double duplex

DNA 3:JP 5

Hetero Amide face duplex

JP 5:DNA 4

Hetero Triazole face duplex

DNA 3:JP 5:DNA 4

Hetero Amide-Triazole double duplex

DNA 5hp :JP 5:DNA 5hp

Hairpin duplex

DNA 6:JP 6
Amide face hetero duplex

JP 6:DNA 4
Triazole face hetero duplex

DNA 6:JP 6:DNA 4
Amide-Triazole face heterodouble duplex

DNA 7$_{hp}$:JP 6:DNA 7$_{hp}$
Hairpin duplex

DNA 8:JP 7
Amide face heteroduplex

JP 7:DNA 9
Triazole face heteroduplex

DNA 8:JP 7:DNA 9
Hetero double duplex

Isothermal titration calorimetry (ITC) experiments with cDNAs titrated against the Janus PNA samples showed good sigmoidal profiles suggesting co-operative binding. The stoichiometry of binding (N) expressed as ratios of number of PNA bases to number of DNA bases involved in binding supported the supported the composition of envisaged duplexes. The enthalpy (ΔH) of duplexes from each face in individual *Janus* PNA complexes are in the range -20.0 to -40.0 kcal/mole. The range of free energy (ΔG) changes for different complexes was in a narrow range of -6 to -9 kcal/mole. The enthalpic contribution to free energy was slightly higher than entropic contribution in almost all complexes, suggesting that the *Janus* PNA:DNA binding reactions are perhaps driven by enthalpic considerations.

DNA 2:JP 1:DNA 1:JP 1:DNA 2

Chapter 5: i-Motif and G-Quadruplex Structures from *Janus* PNA oligomers and their Morphological study by FESEM

Janus PNAs having homo oligomeric C or G sequences on one face can form corresponding i-or G-tetraplexes with stability similar to that from *aeg*-PNA or DNA sequences. The supramolecular assemblies of different *Janus* PNA:DNA, complexes were studied through their morphological features using FESEM. The stability of i-motif C_n tetraplexes depend on the pK_a of N3 of Cytidine determined by pH dependent UV spectroscopy. The thermal stability of *Janus* PNA **JP 1** was maximum at pH 5.1, close to pK_a of N3 of C.

Similarly *Janus* PNA-G_4 also forms conventional G-tetraplexes with the thermal stability that are slightly higher than that of tetraplexes from *aeg*-PNA-G_5 at pH 7.1

i-motif of **JP 1**

G4 of **JP 4**
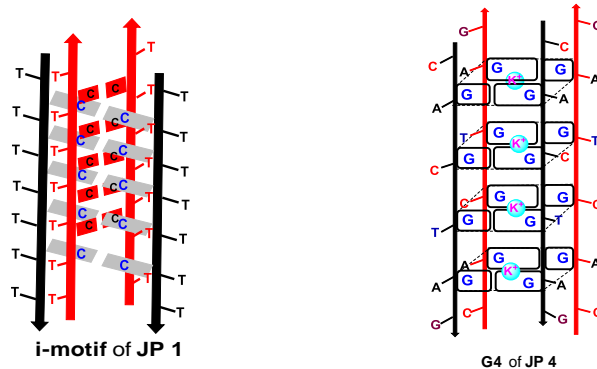
Self-assembly of *Janus* PNA and nanoparticles formation. Supramolecular assembly of different *Janus* PNA oligomers studied by FESEM suggested that *Janus* PNAs, duplexes of triplexes and duplex of duplex show higher order assemblies which result in naofibre formation.
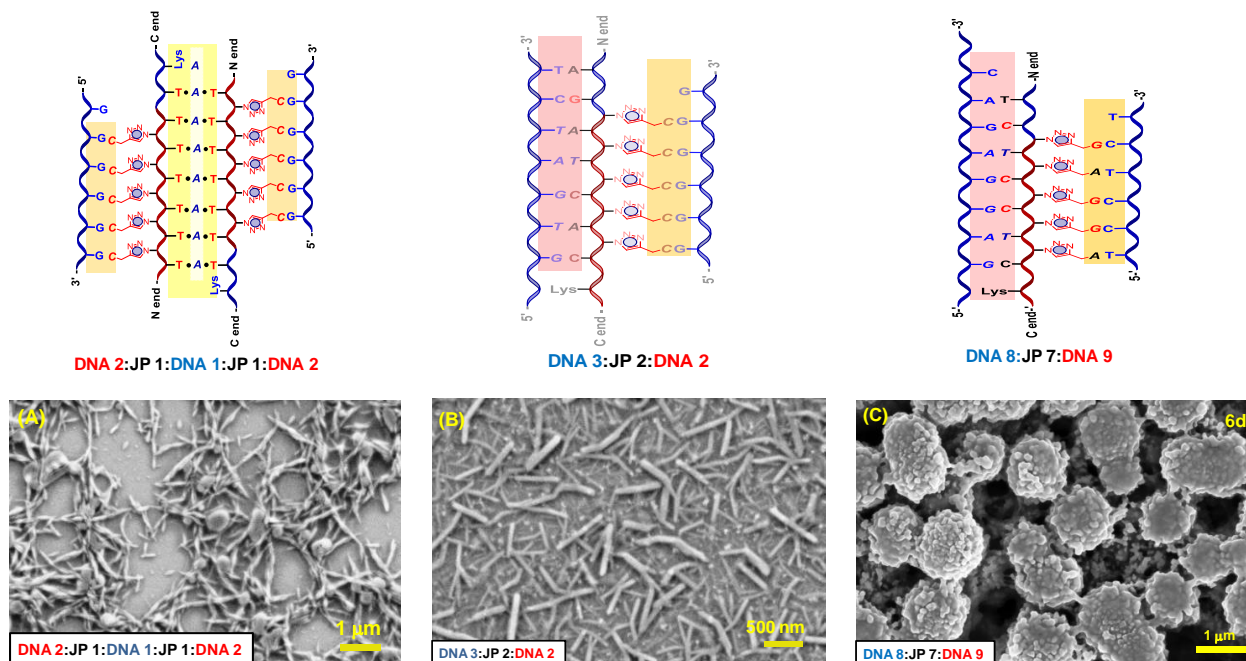
Figure 16. FESEM images of (A) DNA **1**:**JP 1**:DNA **2**. (B) DNA **3**:**JP 2**:DNA **2**. (C) DNA **8**:**JP 7**:DNA **9**.

Summary of Thesis.

- *Janus* PNAs that have nucleobases (A/G/C/T) linked to $C_\alpha(S)$-side chains via triazole link were designed to create new and novel PNA backbones capable of binding to two cDNAs simultaneously on its both faces (amide / triazole). These $C_\alpha(S)$-substituted *aeg*-PNA monomers were assembled to various *Janu*s PNA oligomers by solid phase synthesis, purified by RP-HPLC and characterized by MALDI-TOF spectrometry.

- The thermal stability of duplexes of *Janus* PNA oligomers with each face complementary DNA and double duplexes generated by simultaneous binding of complementary DNAs on both faces were investigated by temperature dependent UV-visible spectroscopy. The T_m of triazole face duplex was always higher than the amide face duplex both in individual duplexes as well as in double duplexes. The *homo Janus* PNA oligomers exhibit double duplex of triplex.

- The duplex from each face (amide / triazole) show characteristic CD bands. Double duplexes of *Janus* PNA can be generated by sequential addition of each face

complementary DNA and the order of addition does not matter. The final double duplex formed has the same conformation as seen by CD.

- The formation of different *Janus* PNA:DNA complexes (duplex, triplex, double duplex and duplex of triplex) were examined by Isothermal Titration Calorimetry (ITC). The results confirmed the stoichiometry of complexes and isuggested that the formation of *Janus* PNA:DNA duplexes are driven by both enthalpic and entropic factors.

- The *Janus* PNAs with ologo C's form terrameric i-motif similar to that of *aeg*-PNA-C_5 *and* DNA C_5 oligomers at acidic pH (< 5.5), while become duplexes. The PNA/*Janus* PNA-G_5 oligomers also form G-quadruplexes. Supramolecular self-assembly of *Janus* PNA:DNA complex were reflected in their FESEM images.

Janus PNAs are a new class of nucleic acid analogues that can simultaneously bind complementary DNA/RNA/PNA strands from two sides by canonical Watson-Crick base pairing and have great potential for new material and biological applications.

Chapter 1

Introduction to *Janus* Peptide Nucleic Acids

1.0 Introduction to nucleic acids

Nucleic acids are the essential biological macromolecules present in all known forms of life. The most supreme biological macromolecules are deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). Their important functions include storage, transmission, and expression of the genetic information within the biological systems (Figure 1.1).
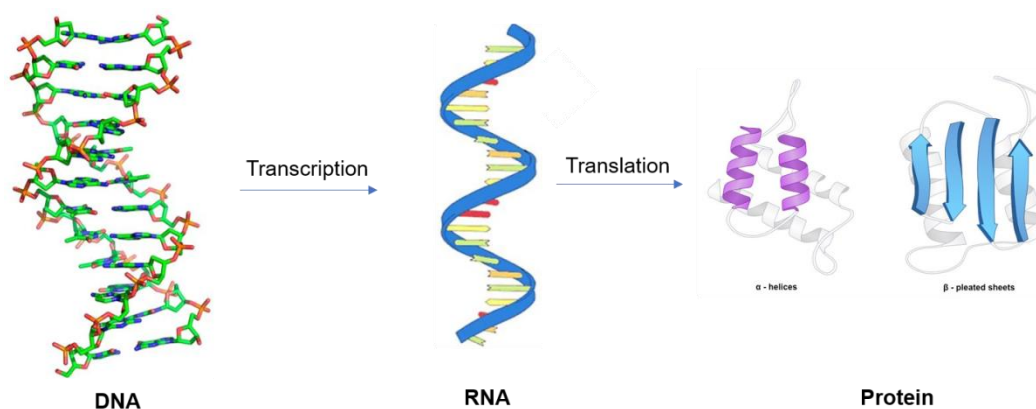
Figure 1.1 Biomacromolecules involved in the flow of genetic information

The DNA and RNA molecules are made up of repeating units of nucleotides. Each nucleotide unit consists of a nitrogenous base (purine or pyrimidine), a pentose sugar[1] and a phosphate group. Both DNA and RNA comprise of four nucleobases: adenine (A) and guanine (G) which are purines, and cytosine (C) and thymine (T, in DNA) or uracil (U, in RNA) that are pyrimidines (Figure 1.2 B). The structures of DNA and RNA differ from each other in their sugar: DNA contains a deoxyribose unit at 2' position, while RNA has a ribose sugar at the same 2' position which has a tremendous effect on stability and conformational preferences of RNA. DNA and RNA also differ in their nucleobase composition. The three heterocyclic bases (A, G and C) are common in both DNA and RNA, but U is present only in RNA and T is found in DNA. In 1953, Watson and Crick proposed that the molecular structure of DNA consists of two helical chains, each coiled around the same axis with a right-handed twist.[2] In these linear copolymers, 3'-5' phosphodiester bonds link successive *β*-D-deoxyribofuranose of the nucleotides to form the negatively charged backbone. The two DNA strands are held together by specific hydrogen bonds between complementary base pairs (A:T & G:C) commonly known as Watson-Crick base pairs to form an antiparallel double helical structure. The backbone of the double helical strands is positioned outside of the main helix with the H-bonded hydrophobic nucleobases stacked perpendicular to the helix axis located in the interior (Figure 1.2).
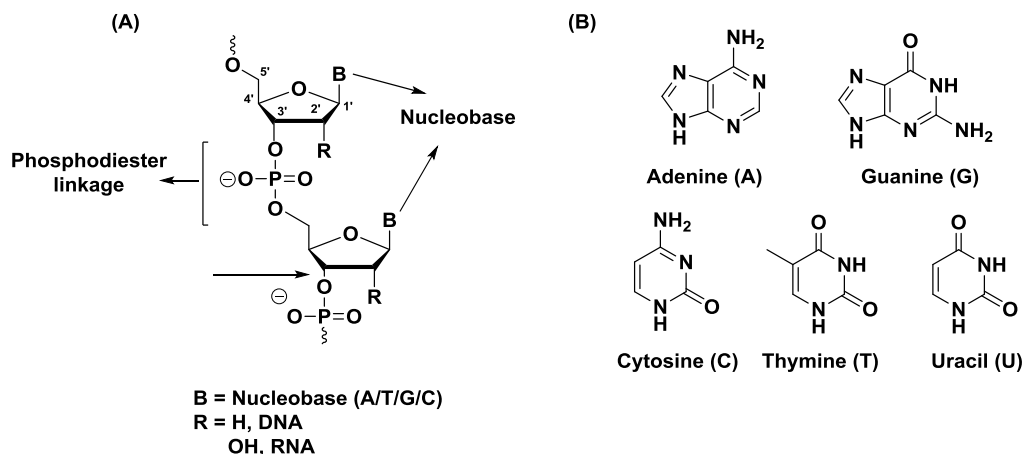
Figure 1.2 Chemical structures of **(A)** DNA and RNA **(B)** Nucleobases

1.1 Base Pairing through Hydrogen bonding

Sequence-dependent molecular recognition between strands in nucleic acids through complementary hydrogen bonding is one of the most important principles of molecular self-assembly that governs information processing in the complex biological systems. The hydrogen bonds are formed specifically between the *amino-keto* tautomer of the bases that establish high fidelity in DNA transcription and translation processes. The N-H groups of the bases are potent hydrogen bond donors, while the sp^2 hybridized electron pairs on the oxygen of the carbonyl (C=O) groups and nitrogens present in the aromatic ring are good hydrogen bond acceptors. This leads to Watson-Crick hydrogen bonding with two hydrogen bonds in A:T base pair and three hydrogen bonds in G:C base pair (Figure 1.3).
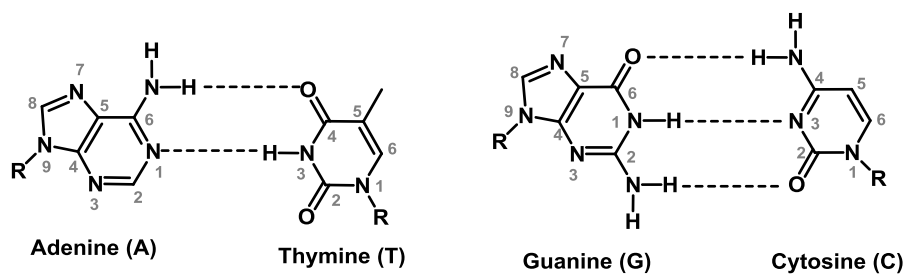
Figure 1.3 Watson-Crick hydrogen bonding for A:T and G:C base pairs[3]

Other important hydrogen bonding pairs are the Hoogsteen[4] (HG) and wobble[5] base pairs. Hoogsteen base pairing is not isomorphous with Watson-Crick base pairing because they have an 80° angle between the glycosidic bonds and 8.6 Å separation of the anomeric carbons (Figure 1.4a, b). Hoogsteen hydrogen bonding has importance in triple helix formation and in protein-

DNA complexes. Wobble base pairing involves non-Watson-Crick pairing between two nucleotides in RNA molecules. Wobble base pairs are fundamental in RNA secondary structure and are important for the proper translation of genetic code (Figure 1.4c, d).
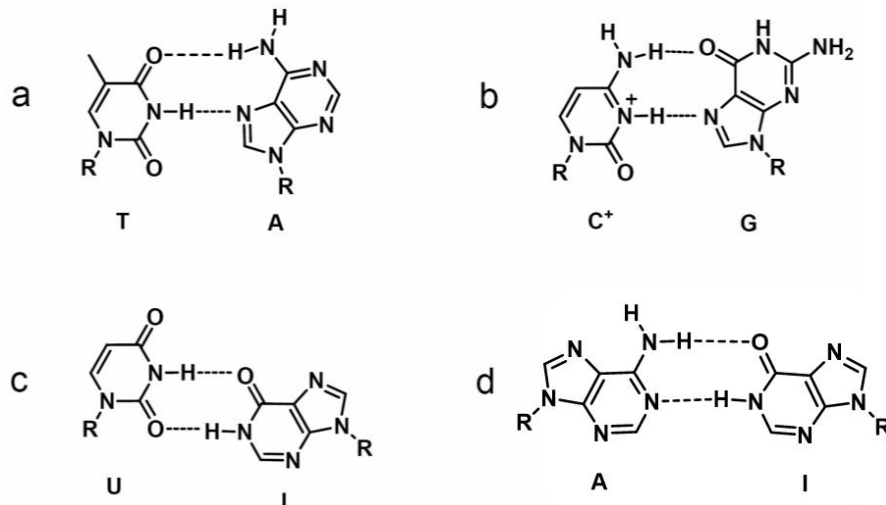
Figure 1.4 (a & b) Hoogsteen base pairing[4] and (c & d) wobble base pairing[5]

1.2 Secondary structures of nucleic acids

DNA exists in various possible conformations like A-DNA, B-DNA, and Z-DNA, although, the most common is the B-DNA which is a right-handed double helix.[6] It has a wide and deep major-groove with a narrow and shallow minor-groove wherein the bases lie perpendicular to the helical axis. A-DNA also forms a right-handed helix where the major groove is deep and narrow while the minor groove is broad and shallow. In both A and B forms of DNA, the Watson-Crick base pairing is maintained along with *anti*-glycosidic conformation. Z-DNA is a left-handed double helix and most favored in alternating G-C sequences. The left-handed helix for Z-DNA is a result of a switch in the glycoside bond with a *syn* conformation (Figure 1.5). The sugar conformation is however different in both forms, with the B-DNA form showing *C2'- endo* puckering and the A-DNA form exhibiting *C3'-endo* sugar-pucker (Figure 1.6).
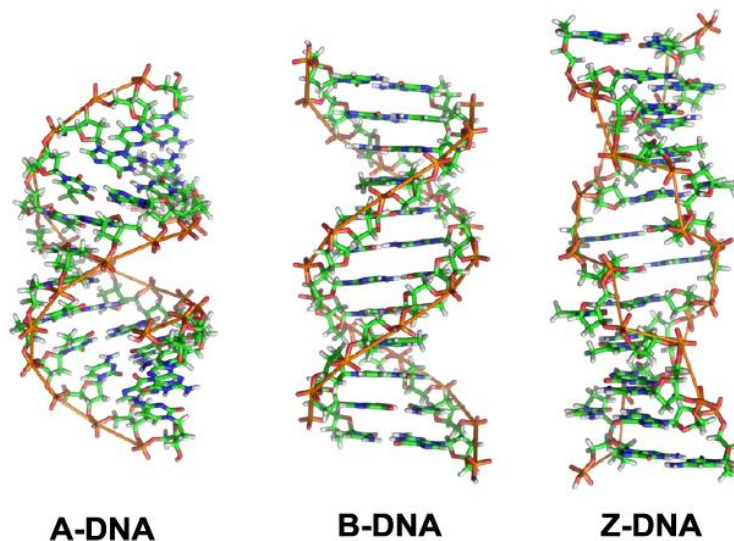
Figure 1.5 Molecular models of A, B and Z-DNA forms of DNA[6]

RNA has greater structural versatility than DNA in the variety of its species, in its diversity of conformations and in its chemical reactivity. The presence of the 2'- hydroxyl group in RNA hinders the formation of a B-form helix and it acquires the A-type helix showing *C3'-endo* sugar puckering. More commonly, RNA is single-stranded, can form complex and unusual shapes such as stem and bubble structures, which occur due to the intrachain base pairing. An example is *t*-RNA, the key molecule involved in the translation of genetic information to proteins.
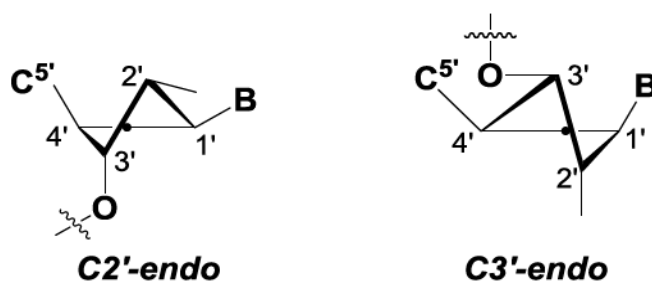
Figure 1.6 Structures of *C2'-endo* and *C3'-endo* sugar puckering[6]

1.3 Applications of nucleic acids

Sequence-specific binding of oligonucleotides to single-stranded RNA or duplex DNA through triple helix formation provides a way to modulate the gene expression. The central feature of biology is the recognition of DNA and RNA sequences by complementary oligonucleotides and it is important for hybridization based biological applications. The study of

such complementary recognition is possible with widely used experimental techniques and diagnostic protocols. The conjugation or attachment of reactive functional groups to oligonucleotides has been investigated to cleave nucleic acids in a sequence specific way which is described as artificial nucleases.[7] The combination of specificity and simplicity makes oligonucleotides highly attractive as diagnostic and research tools.

1.3.1 Oligonucleotide analogs as therapeutic agents

The classical approach of designing drug molecules against protein target requires a deep understanding of 3D structure of the protein, binding site and the binding forces. Since very little is known about the process of protein folding and the lack of generality, this approach of drug discovery has various drawbacks. On the other hand, the nucleotide sequence in RNA and DNA is universal and the understanding of their structures is much better to design drugs for nucleic acid targets. In principle, one can design drugs that are like nucleic acids, are repetitive in its primary structure and bind sequence specifically to the nucleic acid drug targets. In order for the sequence specific recognition to happen, the drug should contain nucleobases which are fundamental units of nucleic acid recognition. In short, small pieces of oligonucleotides containing 15-25 nucleotides can act as a drug which would bind to target nucleic acid and stop further processing of protein production. Two innovative strategies are being tested for inhibiting the production of disease related proteins using such sequence specific DNA fragments as gene expression inhibitors.

The first strategy known as *antisense strategy*[8] aims to selectively impede translation by inhibiting the protein synthesis. Antisense oligonucleotides recognize a complementary sequence on target mRNA through Watson-Crick base pairing and form a duplex which is not recognized for further processing by protein synthesis machinery. Thus, it will retard the expression of corresponding disease related proteins (Figure 1.7b). In the second strategy which is known as *antigene strategy*, the production of unwanted proteins can be stopped by selectively inhibiting the transcription process. In this strategy, oligonucleotides target the major groove of DNA where it winds around the double helical DNA to form triplex involving Hoogsteen hydrogen bonds[9] (Figure 1.7a).

DNA   **mRNA**   **Proteins**

Transcription

Translation

a. Antigene Inhibition   **b. Antisense Inhibition**
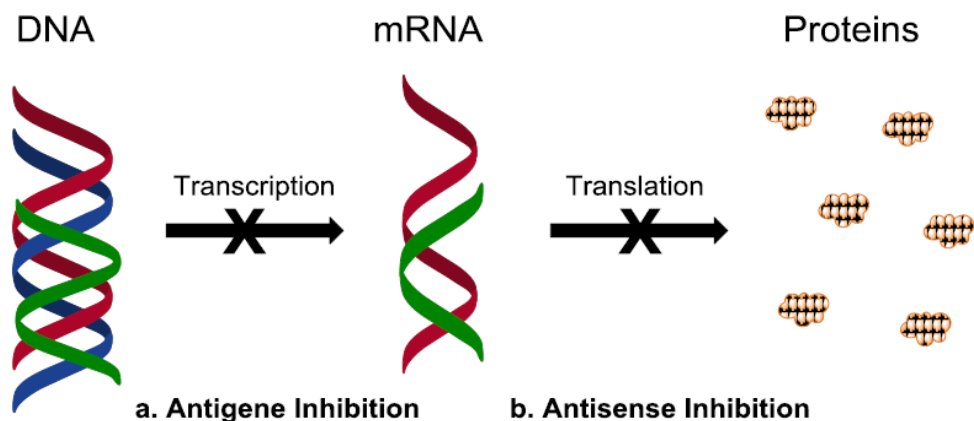
Figure 1.7 Principle of antisense and antigene oligonucleotides[8, 9]

1.4 Antisense oligonucleotides

The principle of antisense drugs is based on the direct utility of sequence information and sequence-specific recognition of nucleic acids and hence such drugs have the potential for the treatment of several viral and bacterial infections, cancerous outgrowths and inflammatory or genetic disorders. Various cellular processes can be inhibited depending on the site at which the antisense oligonucleotide hybridizes to the target nucleic acid. Zamecnik and Stephenson[10] discovered that viral replication in cell culture was inhibited by oligonucleotides (ONs) and these ONs have the potential to act as antisense agents for therapeutics. There are several pre-requisites for oligonucleotides to serve the purpose of inhibiting translation and thereby act as antisense therapeutic agents.[11] These include -

    i) Ease of synthesis in bulk quantities

    ii) *In-vivo* stability to cellular degrading enzymes

    iii) Ability to penetrate the cell membrane

    iv) Retention by the target cells

    v) Ability to interact with their cellular targets (DNA/RNA)

    vi) No non-specific interaction with other macromolecules

Although the requirements for specificity and binding affinity are satisfactorily met by the unmodified oligonucleotides, their susceptibility to hydrolytic enzymes (nucleases) and poor cell permeability are the major obstacles for their use as nucleic acid therapeutic agents. In order to

meet all the above-mentioned requirements, it is necessary for ONs to be chemically modified in a suitable manner.

1.4.1 Chemical modifications of DNA

To improve the rate, affinity, and specificity of oligonucleotide recognition, while enhancing the cell membrane permeability and resistance to nuclease degradation, several chemical modifications of DNA have been attempted[12] in literature (Figure 1.8).
In general, three types of chemical modifications of DNA can be distinguished as

i) Modified phosphate linkages

ii) Modified sugars and

iii) Altered sugar-phosphate backbone.
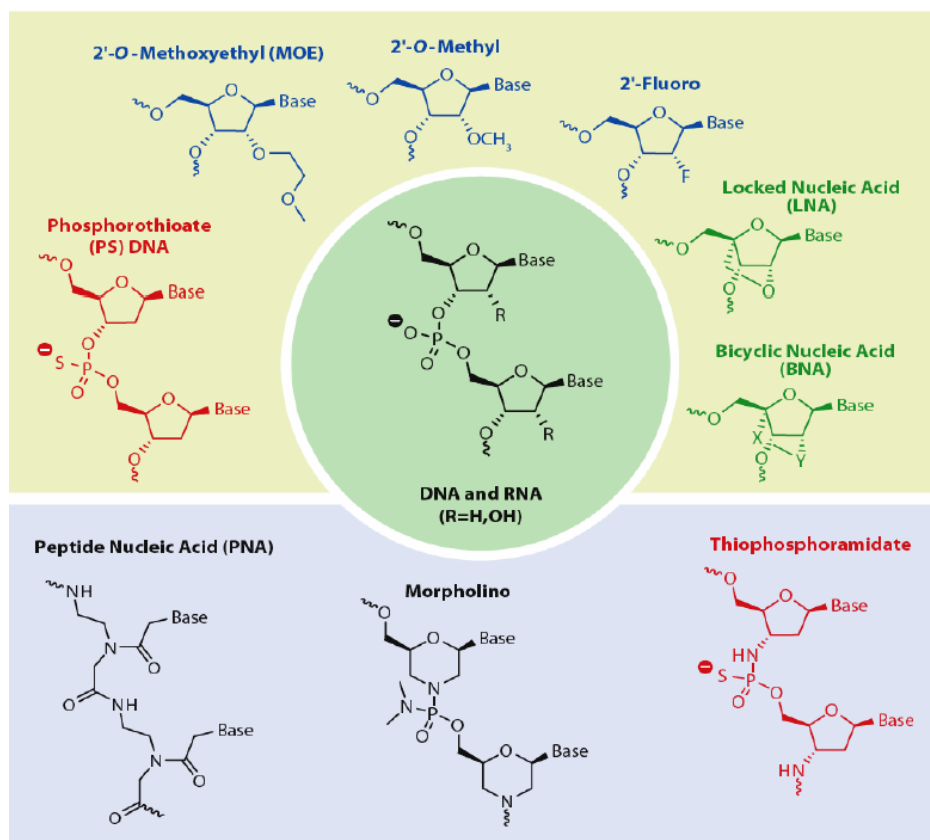
Figure 1.8 Examples of chemical modifications of oligonucleotides[12]

1.4.2 Phosphate backbone analogues

The modifications of phosphate moiety resulted in the development of phosphorothioates[13] **1**, methyl phosphonates[14] **2**, phosphoroamidates[15] **3**, phosphotriesters[16] **4** and

borano phosphonates[17] **5** (Figure 1.9). Phosphorothioate (PS)-containing oligonucleotides differ from natural nucleic acids in that one of the nonbridging phosphate oxygen atoms is replaced with a sulfur atom. Phosphorothioate oligonucleotides are one of the earliest and most widely used backbone modifications for antisense drugs. The substitution of sulfur for oxygen in PS-oligos greatly increases stability to nucleolytic degradation. PS-oligos are able to efficiently elicit the RNase H cleavage of target mRNA, which is critical in the mechanism of action of antisense oligonucleotides.[18] Their binding to plasma proteins protects them from rapid renal excretion and is responsible for increased serum half-life. Vitravene (Fomiversen) is the first FDA approved antisense drug, which is based on PS-oligos. Many antisense drugs which are under various stages of clinical trials incorporate PS-modifications. The binding of PS-oligos to certain proteins, particularly those which interact with polyanions such as hairpin-binding proteins, prove to be their major drawback.[19,20,21]
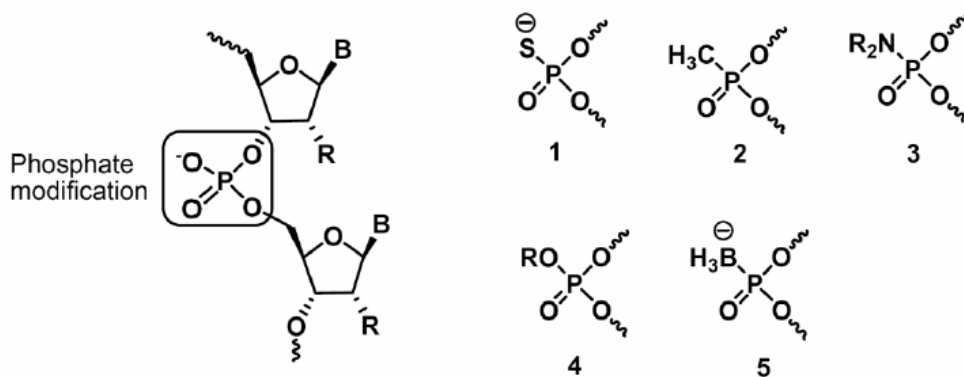
Figure 1.9 Structure of alternative phosphate linkages

Other backbone modifications (methylphosphonates, phosphoroamidates etc.) of oligonucleotides have been less successful at improving the oligonucleotides properties. In methylphosphonates, one of the nonbridging oxygen atoms is replaced with a methyl group and is neutral in charge.[22] Although it provides high nuclease resistance, it does not induce RNase H activity. Additionally, increasing the number of methylphosphonate units in an oligomer leads to loss of affinity towards the target mRNA and to poor water solubility.

1.4.3 Sugar modifications

This class of ONs is represented by nucleotides with alkyl substitutions at the 2'-position of ribose sugar (Figure 1.10 and 1.11). Organization of the sugar into RNA-like *C3'- endo* pucker increases the binding affinity of these ONs towards the complementary RNA.[23]

Furthermore, the 2'-substituent in an oligonucleotide increases the steric bulk near the 3'-phosphate and makes it difficult for degrading enzymes to cleave the phosphodiester bond. The increase in binding affinity of 2'-modified ONs is energetically driven by the electronegative substituent at the 2'-position. Among all 2'- modified ONs, 2'-fluoro modification (Figure 1.10a) imparts the highest binding affinity towards the target RNA. 2'-O-Methyl (Figure 1.10b) group enhances the binding affinity to a lesser extent than do the 2'-fluoro modification but impart a substantial degree of nuclease resistance to the corresponding oligonucleotide. 2'-O-Methoxyethyl[24] (MOE) (Figure 1.10c) is currently the most advanced in the 2'-modified series and many antisense drugs having this modification have entered clinical trials. Unfortunately, the lack of RNase H activity restricts the use of 2'-modified oligonucleotides for antisense purpose. To overcome this drawback, the gapmer strategy has been used where regions of 2'-modified residues flank a central DNA region of the oligonucleotide.[22]

a) 2'-Fluoro   b) 2'-O-methyl   c) 2'-O-methoxyethyl
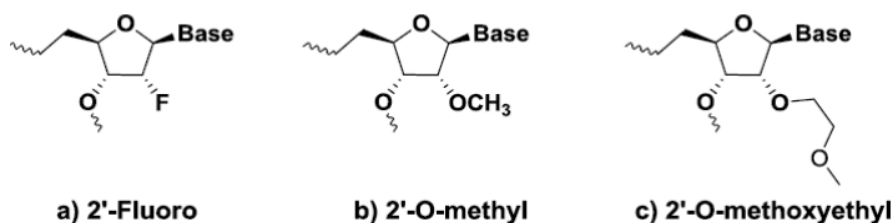
Figure 1.10 2'-modified oligonucleotides

1.4.3a Locked nucleic acid (LNA). These oligonucleotides developed by Jesper Wengel *et al.*[25] in 1998 are one of the most promising class of chemically modified ONs. These analogs are bicyclic systems that contain a methylene bridge that connects the 2'-O- of the ribose with the 4'-C, locking the ribose in *C3'-endo* conformation (Figure 1.11). Introduction of LNA into a DNA sequence induces a conformational change of the DNA:RNA duplex towards the A-type helix[26] but prevents the RNase H cleavage of the target RNA. LNAs and LNA:DNA chimeras have been shown to efficiently inhibit gene expression when targeted to a variety of regions within the luciferase mRNA.[27] LNA shows remarkably increased hybridization properties relative to a DNA:RNA duplex and improves nuclease resistance. Few analogs of LNA have been reported, which have improved activity and/or toxicity profiles in animals.[28,29] Thus, LNA offers attractive properties, such as nuclease stability, high target affinity, potent biological activity and lack of acute toxicity.
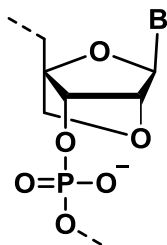
Figure 1.11 Locked nucleic acid (LNA)

1.4.4 Sugar-phosphate backbone modifications

In addition to phosphodiester and sugar modifications, replacement of the sugar phosphate backbone with isosteric structures has been devised. The concept of conformational restriction has been used widely to enhance binding affinity and bio stability. Some of the DNA and RNA analogs developed with modified sugar phosphate backbones are described below.

1.4.4a Morpholino oligonucleotides (MF). Phosphorodiamidate morpholino oligonucleotide has morpholine ring replacing the furanose ring in DNA/RNA. It has the phosphorodiamidate linkage which connects the morpholine nitrogen atom with the hydroxyl group of the 3'-side residue (Figure 1.12).[22] Morpholino ONs are non-ionic, and therefore unlikely to have unwanted electrostatic interactions with nucleic acid binding proteins. These ONs are stable to nucleases and have a similar affinity as in DNA:DNA duplexes. However, morpholino ONs do not activate the RNase H and are primarily used in translation arrest or in other steric blocking mechanisms, such as alteration of splicing.[30,31] Eteplirsen (EXONDYS 51) is the first morpholino oligo-based antisense drug approved by FDA for the treatment of Duchenne muscular dystrophy (DMD).
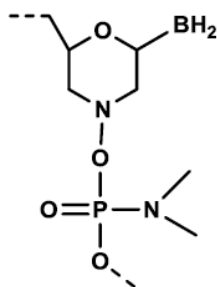
Figure 1.12 Morpholino oligonucleotide (MF)

1.4.4b N3'-P5' phosphoroamidates (NPs). N3'-P5' phosphoroamidate (NPs) is a modified phosphate backbone, in which the 3'-oxygen of the deoxyribose ring is substituted with an amino

group (Figure 1.13). NPs are resistant to nucleases and have high affinity towards a complementary RNA strand but do not activate the RNase H.[32] The sequence specificity of phosphoroamidate-mediated antisense effects by steric blocking of translation initiation has been demonstrated in cell culture and *in vivo*.[33]
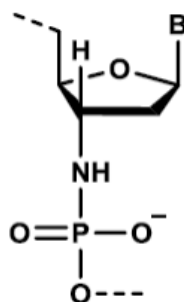
Figure 1.13 *N*3'-*P*5' phosphoroamidates (NPs)

1.5 Peptide nucleic acids (PNA)

Peptide nucleic acids (PNAs) are a radically different class of oligonucleotide modifications that contain a peptide link instead of phosphate link. PNA was first introduced by Peter Nielsen and coworkers in 1991.[34] PNAs are highly resistant to degrading enzymes like proteases and nucleases, exhibit high binding affinity towards target DNA/RNA, but do not activate RNase H and, as such, have been used primarily in translation inhibition and splicing modulation antisense mechanisms. PNAs seems to be non-toxic, as they are uncharged molecules with low affinity for proteins that normally bind nucleic acids.

PNAs are DNA analogs where the sugar-phosphate backbone is replaced with a pseudopeptide backbone in the form of 2-aminoethyl-glycine linkage. Nucleobases are attached through a methylene carbonyl linker to this backbone at the amino nitrogen. The PNA backbone is constituted by six atoms for each repeating unit and a two-atom spacer between the backbone and the nucleobase, similar to the natural DNA (Figure 1.14).[35] PNA was originally designed and developed as a mimic of a DNA recognizing, major groove binding, triplex-forming oligonucleotide.[35,36] However, the polyamide backbone of PNA has proven to be a surprisingly good structural mimic of the sugar-phosphate backbone of nucleic acids.
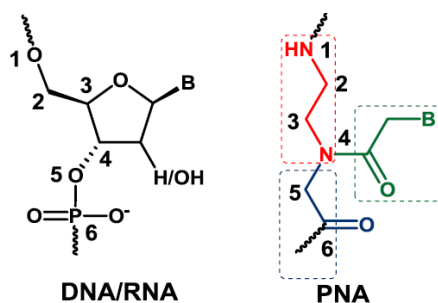
Figure 1.14 Chemical structures of DNA/RNA and PNA

The advantages of PNA over the conventional antisense oligonucleotides are numerous, partially due to the high flexibility and the absence of charge in the backbone. PNAs being neither peptide nor nucleic acids are resistant to both proteases and nucleases and consequently have a longer life span in the cellular environment. PNA hybridizes with complementary DNA/RNA with thermal stability superior to DNA:DNA or DNA:RNA duplex. This results from a decrease in electrostatic repulsion between the two negatively charged strands in DNA/RNA duplexes. They have higher mismatch discrimination and form selective duplexes upon binding to complementary DNA or RNA sequences. Therefore, PNA has attracted wide attention in medicinal chemistry for the development of gene therapeutics, especially antisense or antigene drugs.

1.5.1 PNA structure

PNA binds to a complementary DNA/RNA through classical Watson-crick base-pairing mechanism. The PNA bases form a helical π-stack similar to DNA but the smaller twist of the PNA double helix and the larger π-overlap between the neighboring bases makes it different from the DNA. NMR methods and X-ray crystallography have been used to derive the three-dimensional structures of the major families of PNA complexes including PNA:RNA,[37] PNA:PNA,[38] PNA:DNA duplexes[39] and PNA$_2$:DNA triplex[40] (Figure 1.15). Although PNA oligomer, to some extent is able to structurally adapt to the oligonucleotide complement, it also has a preferred structure of its own, termed the '*P-form*' helix.
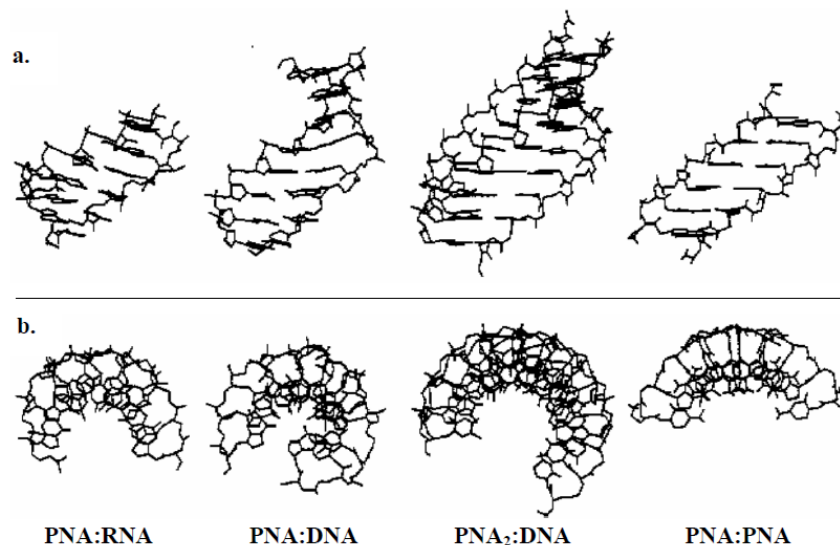
Figure 1.15 Structures of various PNA complexes shown in a) side view and b) top view[41]

The *P-form* helix constitutes a very wide PNA duplex (28 Å diameter) with an accordingly large base pair helical displacement and a very large pitch (18 bp). A canonical *B-form* helix which is typical for DNA duplexes has a diameter of 20 Å and a pitch of 10 bp per turn. A canonical *A-form* helix, typical of RNA duplexes, also has a diameter of 20 Å but a pitch of 11 bp per turn, and the base pairs are tilted 20° relative to the helix axis. Also, the base pairs are displaced away from the helix leaving a central 'tunnel' in the helix, analogous to that seen in the *P-form*. These structures suggest that PNA can adapt well to its nucleic acid partner, as the RNA strand in the PNA:RNA duplex is essentially in A-form conformation, whereas PNA:DNA duplex adopts a B-form conformation.

1.5.2 Physico-chemical properties of PNA

PNA has proved itself a promising antisense or antigene agent on the basis of its superior properties, such as highly sequence-specific binding to the complementary DNA/RNA targets, high biological and chemical stability, high mismatch discrimination etc.

*1.5.2a **Duplex formation with complementary oligonucleotides***. PNA was originally designed for sequence-specific targeting of double-stranded DNA via major groove recognition. PNA hybridizes to complementary oligonucleotides obeying Watson-Crick base pairing rule. In DNA:DNA duplexes, the two strands are always in antiparallel orientation (with the 5'-end of one strand opposed to the 3'-end of the other). However PNA:DNA adducts can be formed in two different orientations, arbitrarily termed *parallel* and *antiparallel* (Figure 1.16). Both

14

adducts are formed at room temperature, with the antiparallel orientation showing higher stability.[42] This creates the possibility for PNAs to bind two DNA tracts of opposite sequence.

Figure 1.16 Antiparallel and parallel modes of PNA:DNA duplex formation

1.5.2b Triple helix formation of PNA. Polypyrimidine PNAs are able to form stable adducts with complementary polypurine DNA, through the formation of PNA_2:DNA triplexes.[43] The base pairing in these complexes occurs via Watson-Crick and Hoogsteen hydrogen bonds. When only one PNA strand is used to form a PNA_2:DNA triplex, both strands are necessarily either antiparallel or parallel to DNA strand. When two different homopyrimidine PNA sequences are used, Watson-Crick PNA strand is oriented in antiparallel and the Hoogsteen strand is in parallel orientation to form a stable triplex with homopurine strand of DNA. The sequence specificity of triple helix formation is based on the selectivity of formation of the intermediate PNA:DNA duplex, whereas binding of the third strand contributes only slightly to selectivity. The stability of structures enables PNA to perform strand invasion,[43,44] a property uniquely shown by PNAs.

1.5.2c Strand invasion by PNA. The unique property of PNAs to displace one strand of DNA double helix to form strand invasion complexes,[45] is a favorable attribute for their application as antigene agents. Four modes of binding for sequence-specific targeting of double-stranded DNA by PNA have been identified (Figure 1.17), three of which involve invasion of the DNA duplex by PNA strands.
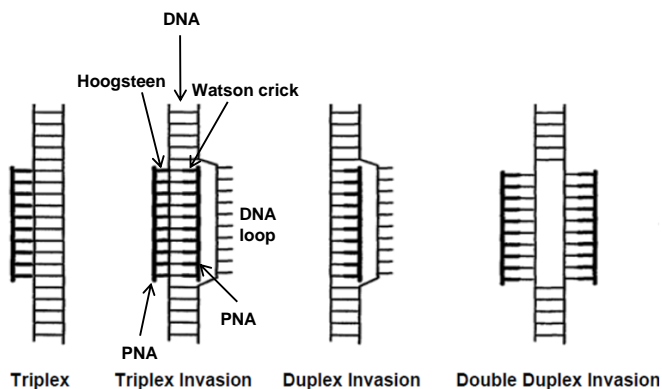
Figure 1.17 Various modes for binding of PNA to sequence specific targets in *ds*DNA[46]

It is possible for PNA (homopurine) single strand to either invade (duplex invasion) via Watson-Crick base pairing, or alternatively, invasion may be accomplished by two pseudo-complementary PNA strands, each of which binds to one of the DNA strands of the target (double duplex invasion). These pseudo complementary PNAs contain modified adenine and thymine nucleobases[46] (Figure 1.18) that do not allow stable hybridization between the two complementary sequence PNAs, but does permit good binding to the DNA. The 'triplex invasion' requires a homopurine DNA target and complementary homopyrimidine PNAs that bind the purine DNA strand through combined Watson-Crick and Hoogsteen base paring *via* formation of a very stable PNA2:DNA triplex. For many applications, the two PNA strands are connected in a bis-PNA designed such that the one strand is antiparallel (W-C strand) and the other strand is parallel (H-strand) to the DNA target.
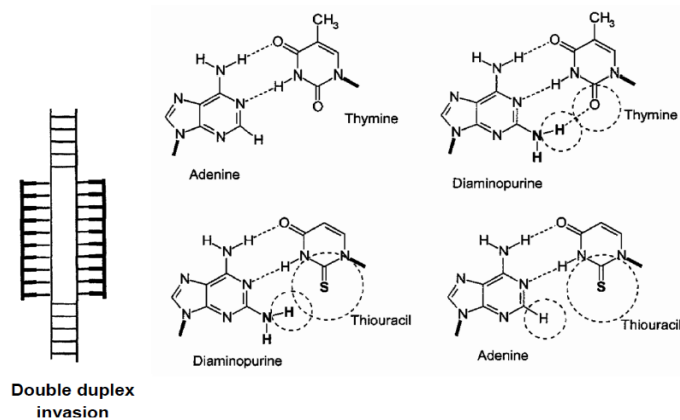
Figure 1.18 Double duplex invasion of peudocomplementary PNAs[46]

1.5.2d G-Quadruplex and i-motif formation by PNA. DNA and RNA oligomers that contain multiple stretches of consecutive guanine (G) nucleotides are able to fold into a stable secondary structure known as G-quadruplex[47,48] that is gaining increasing attention due to its implication in regulation of gene expression.[49] Balasubramanian *et al.*[50] have shown the formation of quadruplex composed entirely of PNA (Q-PNA) (Figure 1.19b). A homologous PNA (i.e. a PNA having the same sequence as the target) forms a stable PNA2:DNA2/RNA2 hybrid quadruplex by disrupting a bimolecular DNA/RNA G-quadruplex (Figure 1.19a).[51] Ganesh *et al.*[52] demonstrated the formation of C-C$^+$ tetraplex (i-motif) with unmodified PNAs TC$_4$ and TC$_8$ in acidic pH. The tetraplex formation was monitored using UV-thermal melting at 295 nm, show a reverse sigmoidal pattern, which is characteristic of C-C$^+$ tetraplex formation (Figure 1.19c).
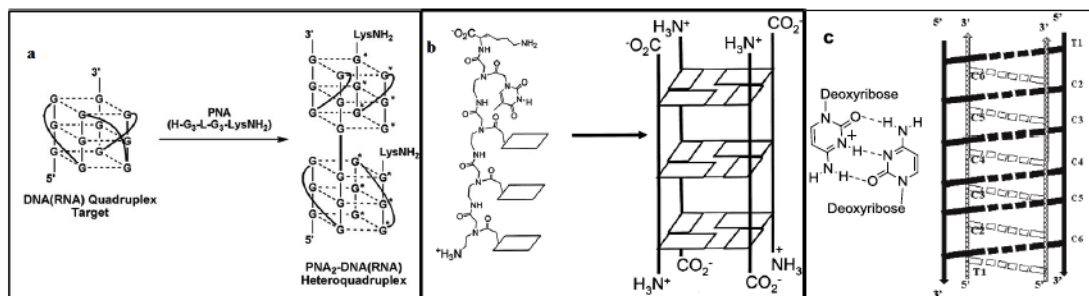
Figure 1.19 **(a)** Schematic drawing for PNA:DNA/RNA homologous quadruplex[51] **(b)** Schematic diagram for PNA-Q4 quadruplex[50] **(c)** Schematic diagram of i-motif formation[52]

1.5.2e Higher specificity towards target nucleic acid. PNA shows greater specificity in binding to complementary DNA. A PNA/DNA mismatch is more destabilizing than a mismatch in a DNA/DNA duplex. A single mismatch in mixed PNA:DNA (15-mer) sequence, destabilizes the duplex by 8-20 °C (15 °C on average), while in the corresponding DNA:DNA duplex, a single mismatch destabilizes the duplex by 4-16 °C (11 °C on average).[53]

1.5.2f Stronger binding independent of salt concentration. Another important consequence of the neutral backbone is that the T_m values of PNA:DNA duplexes are practically independent of salt concentration. In contrast, the T_m values of DNA:DNA duplexes are highly dependent on ionic strength.[54]

1.5.2g Solubility of PNA. PNAs are charge-neutral compounds and hence have poor water solubility compared with DNA. Neutral PNA molecules have a tendency to aggregate to a degree that is dependent on the sequence of the oligomer. PNA solubility is also related to the length of the oligomer and to the purine/pyrimidine ratio.[54,55] Some of the recent modifications, including the incorporation of positively charged lysine residues (carboxy-terminal or backbone modification in place of glycine), have shown improvements in solubility of PNA.

1.5.2h Cellular uptake of PNA. Although PNA binds to complementary DNA/RNA with high affinity, specificity and stability in biological fluids, the progress in the exploration of PNA as antisense/antigene agents and gene expression regulation has been hampered by their poor cellular uptake. Thus, efficient cellular delivery systems for PNAs are required if these are to be developed into antisense and antigene agents. However, a number of transfection protocols for PNA have been established like microinjection, electroporation, co-transfection with DNA, conjugation to lipophilic moieties, conjugation to cell penetrating peptides etc.[56] To address the

issues like poor cell penetration, solubility and ambiguity in binding orientation, various chemical modifications of PNA as describe below.

1.5.3 Chemical modifications of PNA

The antisense antigene potential of PNAs can be improved by enhancing the binding affinity for DNA and RNA by suitable conformational preorganization. One approach for improving DNA binding affinity is the design and synthesis of preorganized PNAs that prefer a right-handed helical conformation which can be achieved by adding of substituents to the backbone or by cyclization of the PNA backbone (Figure 1.20).[57,58]
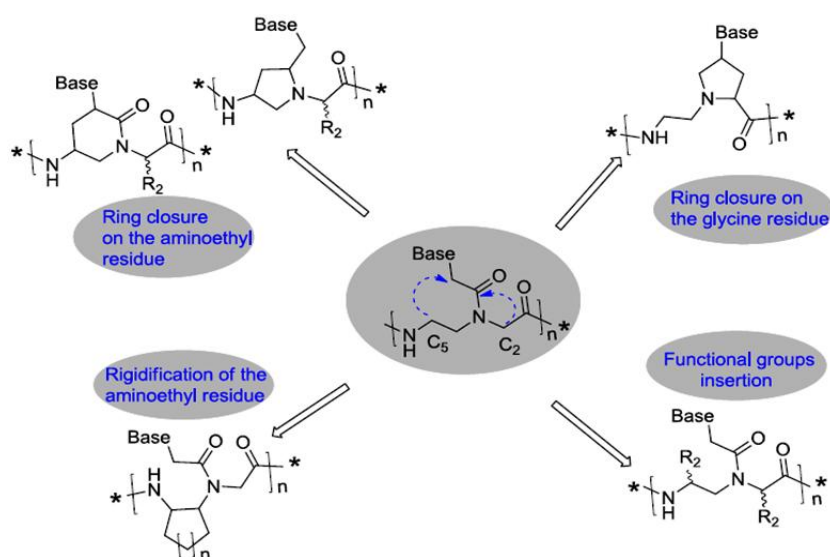
Figure 1.20 Strategies for inducing preorganization in the PNA structure

1.5.3a Preorganization of acyclic PNAs. The N-(2-aminoethyl) glycine unit in PNA backbone is a versatile motif for the modification of the PNA. Replacement of glycine by other α-substituted amino acids and substitution in ethylenediamine results in chiral PNAs that bear substituents at different positions (α, β, γ) (Figure 1.21)
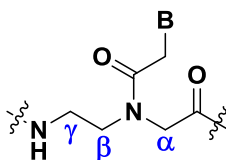
Figure 1.21 Structures of achiral PNA, B = A/T/G/C. (A = Adenine, T = Thymine, G = Guanine and C = Cytosine)

C$_\alpha$-PNAs: The first C$_\alpha$-PNAs were reported by Nielsen *et al.*[59] in 1994, derived from both L and D alanine (Figure 1.22a). PNA:DNA duplex stability derived from D-alanine in PNA is almost equal to unmodified PNA whereas C$_\alpha$-substituted PNA derived from L-alanine drastically destabilised the duplex stability. PNAs derived from D-lysine (Figure 1.22b) in place of glycine exhibited higher T_m than that from the corresponding PNA derived from L-lysine, with T_m more than that with unmodified PNAs.[60] C$_\alpha$-substituted PNAs derived from D-amino acids prefer P-helix formation and mismatch discrimination ability is more than those of unmodified PNAs.[61] PNA oligomer containing three units of Lys-D-PNA and its peptide conjugate tested for cellular uptake ability, were found to be taken in most of the tested cells.[62]

D-Arginine in place of glycine leads to α-guanidinium PNAs (GPNAs, Figure 1.22c) that showed destabilization of the derived PNA:DNA duplex. However, incorporation of multiple α-guanidinium PNA units (α-GPNA) in a PNA decamer improved the binding affinity, partly from electrostatic interactions between guanidinium group and phosphates.[63] C$_\alpha$-PNAs with negatively charged side chains (derived from glutamic acid and aspartic acid) destabilized the PNA:DNA duplex due to electrostatic repulsions between negatively charged PNA backbone and polyanionic DNA backbone.[64]

C$_\alpha$-substituted PNAs derived from α-amino acids having bulky side chains such as tyrosine, histidine, tryptophan, phenylalanine and valine showed lower T_m values due to steric hindrance.[65] 1-(2,2-dimethylcyclobutyl) ethanone containing α-PNA (Figure 1.22e) showed almost equal binding affinity with DNA and RNA.[66] Gem-dimethyl substitution at C$_\alpha$ increases the T_m of derived PNA:DNA duplexes,[67] and the homologous aminopropyl-(α,α-dimethyl)glycine (*apdmg*) PNAs improved the binding affinity to DNA. These gem-dimethyl substituted PNAs have showed preferential binding to DNA than to RNA.

C$_\beta$-PNAs: In 2011, Sugiyama *et al.*[68] reported the first C$_\beta$-substituted PNA bearing a methyl group at β-position using D and L enantiomers of alanine (Figure 1.22f). C$_\beta$-PNA oligomer derived from L-alanine showed similar binding affinity as that of unmodified *aeg* PNAs, whereas the corresponding C$_\beta$-PNA derived from D-alanine did not bind to DNA. Circular dichroism revealed that C$_\beta$-*S*-PNA exhibited a right-handed helical structure whereas C$_\beta$-*R*-PNA exhibited left-handed helical structure.[68]
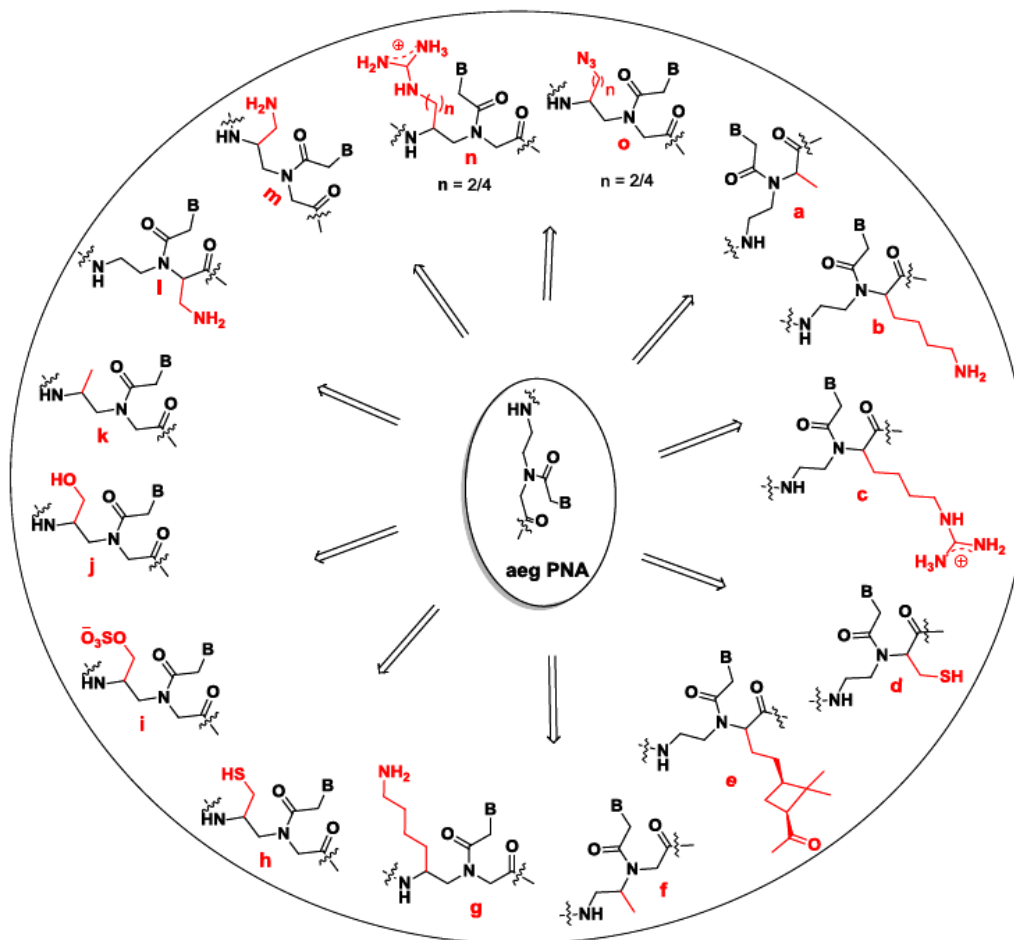
Figure 1.22 Modified acyclic PNAs.

C_γ-**PNAs:** Although the first chiral γ-PNA was reported in 1994, γ-PNA oligomers did not appear in literature until 2005.[69-73] C_γ-(L)-Lys-modified PNAs (Figure 1.22g) stabilized PNA:DNA duplexes and showed good mismatch discrimination relative to unmodified PNAs, whereas γ-D-Lys-PNAs markedly destabilized the duplex.[74] Cysteine based PNA substitution (Figure 1.22h) at the N-terminus of a PNA oligomer can be used for the native chemical ligation with PNAs carrying thioesters to yield long chain PNAs. The effect of stereochemistry on DNA binding was more effective in C_γ-PNAs than that of C_α-PNAs.[75] Homopyrimidines incorporated with C_γ-PNA bearing a sulfate group (Figure 1.22i) formed a triplex with lower stability which arises from the repulsions between negatively charged phosphate DNA and sulfate of modified PNAs. The modified PNA could be lipofected into human breast cancer (SKB3) cells and exhibited antigene activity against ErbBr gene.[76]

Serine or alanine based γ-PNAs (Figure 1.22j and 1.22k) revealed that single-stranded γ-backbone modifications preorganized to right-handed helix, similar to that in PNA:DNA duplex.[77,78] Ly *et al*.[79] have reported a variety of γ-PNAs and demonstrated that C_γ-position can accommodate hindered side chains unlike C_α-PNAs without affecting the hybridization properties considerably. The cellular uptake of a fully modified alternate C_γ-GPNA decamer by HeLa cells is comparable to that of TAT transduction domain. Recently, Manicardi *et al*.[80] reported the inhibition of micro-RNA by GPNAs carrying an arginine side chain. Anti-miR-210 activity of PNAs in leukemic K562 cells was examined using a series of modified 18-mer PNAs (α or γ, PNAs conjugated with arginine octamer and unmodified PNAs). The best anti-miR-210 activity was observed with γ-GPNAs with consecutive placement.[80] Ly *et al*.[81] have reported C_γ (*R*) that ethylene glycol derived from L-serine PNA ($^{R-MP}$γPNA), have tendency to preorganize into a right-handed helix and hybridize to DNA and RNA with greater affinity and sequence selectivity. They were also more water soluble and less aggregating than the parental PNA oligomers, and these PNAs ($^{R-MP}$γPNA) also efficiently suppressed *in-vitro* translation.[82]

Ganesh *et al*.[83] have reported the design and synthesis of chiral C_α- and C_γ- aminoethyl PNAs (*am*-PNAs) with substitutions in the PNA backbone (Figure 1.19*l* and 1.19m). The *am*-PNAs formed stable PNA:DNA duplexes and the order of stabilization was, γ-(*S*)- *am*-PNA > α-(*R*)- *am*-PNA > α-(*S*)-*am*-PNA. The *am*-PNAs are taken up by HeLa cells, with decreasing order of uptake efficiency as γ-(*S*)-*am*-PNA > α-(*R*)-*am*-PNA > α-(*S*)-*am*- PNA. Recently[82] it was shown in γ-GPNAs (Figure 1.19n, n = 2) and γ-azido PNAs (Figure 1.19o) that side chain lengths is important for hybridization with DNA. The guanidinium PNAs with shorter spacer chain increased the PNA:DNA duplex stability. These PNAs taken up in 3T3 and HeLa cells were visualized by confocal microscopy and quantified using fluorescence assisted cell sorter (FACS).[84] The azido functionality of C_γ-substituted methylene/butylene azido PNAs enable the attachment of multisite labeling and introduction of fluorophores in a single step through click reaction without any protection/deprotection steps. The azido fluorescent PNA oligomers have been shown to accumulate around the nuclear membrane in 3T3 cells.[85]

1.5.3b Conformational preorganization through cyclic PNAs. The PNA being acyclic, is conformationally flexible and formation of ordered PNA:DNA/RNA complexes is accompanied by enthalpic advantage through hydrogen bonding and base stacking interactions but has an undesirable loss in entropy of the PNA strand.[86] A high rotation barrier is encountered in the

interconversion of *cis* and *trans* rotamers around the tertiary amide linkage in PNA and leads to different PNA:DNA/RNA hybridization kinetics in parallel and antiparallel hybrids.[87,88] Ganesh *et al*.[89] have synthesized many modified cyclic PNAs that can potentially preorganize PNA to form PNA:DNA/RNA complexes and achieved hybridization preferences (decreased loss in entropy) for either DNA or RNA complementation and parallel or antiparallel orientation.

Many strategies have been developed to enrich populations of single-stranded PNAs that have the favorable, pre-organised conformation for binding to complementary DNA/RNA. The preorganized conformers could trigger a shift in equilibrium toward the desired complexation because of the net reduced entropy loss upon complexation. The strategies to conformational preorganization are based on the introduction of methylene/ethylene groups to bridge the aminoethyl-glycyl backbone and methylene carbonyl side chain to generate diverse five- or six-membered nitrogen heterocyclic analogs. Some of these strategies are described below in detail.

Aminoprolyl PNA (ap-PNA): 4(*R/S*)-Aminoprolyl PNA synthesized by the introduction of a methylene bridge between β-carbon atom of the aminoethyl segment and the α''-carbon of the glycine segment on *aeg*-PNA backbone (Figure 1.23).[90] None of the homochiral aminoprolyl thyminyl PNAs corresponding to any of the diastereomers bound to target DNA sequences[91] and this might be attributed to high rigidity in the backbone resulting in structural incompatibility.[92]

Aminoethylprolyl PNA (aep-PNA): Chimeric aminoethylprolyl PNA (*aep*-PNA) analogues arise from connecting the glycine on backbone with side chain to give 4(*S*)-2(*S/R*)-*aep*-T oligomers (Figure 1.23) that showed strong and specific binding properties toward target DNA sequences.[93] The homooligomer *aep*-PNA-T$_8$ with (2*R*,4*R*) stereochemistry displayed significant stabilization of the complexes with poly rA,[94] while the adenine A$_8$ homooligomer with (2*S*,4*S*) stereochemistry showed improved binding to the target DNA.[95] Nucleobase dependent binding and orientation selectivity were observed with these analysis.

Aminoethylpyrrolidinone PNA (aepone-PNA): The aminoethylprolyl-5-one thymine (*aepone*) monomers contain cyclic amide and were incorporated into *aeg*-PNA-T$_8$ backbone at different positions.[96] The *aepone*-PNAs (Figure 1.23) showed remarkable stabilization of derived PNA$_2$:DNA triplexes compared to *aeg*-PNA.[97] The homooligomers of *aepone*-PNA-T$_8$ bound to the complementary 12-mer DNA more strongly than that of cRNA. This suggested that structurally preorganized PNAs can discriminate their binding selectivity to DNA or RNA.
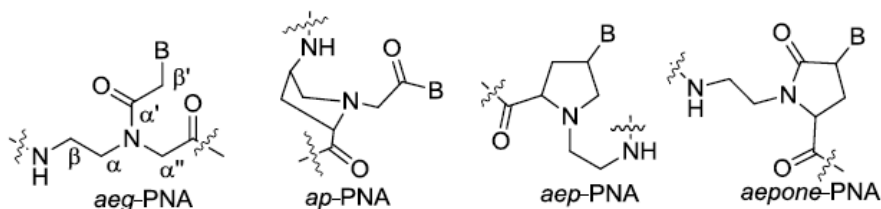
Figure 1.23 Cyclic PNAs: *ap*-PNA, *aep*-PNA and *aepone*-PNA

trans-Cyclopentane PNA (tcypPNA): Daniel H. Appella *et al.*[98] incorporation of (*S,S*)-*trans*-cyclopentane (*t*cyp) units into the PNA backbone improves PNA-DNA duplexes, triplexes, and quadruplexes stability.
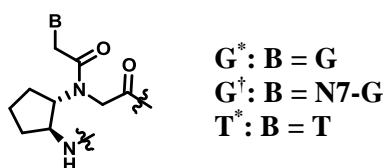
Figure 1.24 (*S,S*) *tcyp*PNA

Pyrrolidine PNAs: Deletion of the *endo*cyclic carbonyl group in pyrrolidinone PNA gave the pyrrolidine PNA[99] (Figure 1.24a). The derived (2*R*,4*S*) stereomeric homo adenylate chimeric *aeg*-PNA oligomer formed a stable complex with both DNA and RNA.[100] The (2*R*,4*R*) version of this PNA analogue was shown to bind target DNA and RNA with high affinity and kinetic selectivity toward RNA.[101] (2*S*,4*S*) pyrrolidine PNA analogue, and its thyminyl oligomer resulted in a decreased binding efficiency with target DNA and RNA sequences.[99] PNA:DNA dimer prepared from (2*R*,4*R*) pyrrolidine PNA-T, which when placed in a PNA-DNA chimera lead to decreased DNA triplex stability.[102]

Introduction of the α'-β-methylene bridge led to another pyrrolidine-PNA (Figure 1.24b)[103] having the nucleobase away from the pyrrolidine ring by one carbon. The (2*R*,4*S*) pyrrolidine-T monomer when introduced into the middle of the *aeg*-PNA-T_8, bound to the target DNA better than the diastereomeric (2*S*,4*S*) PNA. This is an example of stereochemical discrimination effects in PNA:DNA recognition. The 4*R* pyrrolidine-T PNA sequences stabilized the PNA:DNA duplex with a significant difference in parallel/antiparallel binding compared to *aeg*-PNA:DNA duplexes,[104] while the PNAs with 4*S* modifications destabilized the duplexes without much parallel/antiparallel binding differences as compared to the unmodified hybrids. The enantiomeric pairs (2*S*,4*S*) and (2*R*,4*R*) formed antiparallel complexes with RNA much stronger than that of *aeg*-PNA or other diastereomers.
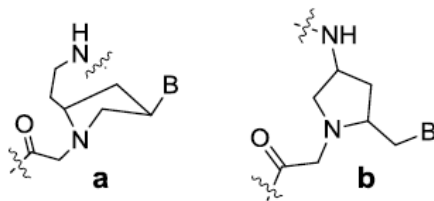
Figure 1.24 Pyrrolidine PNAs[96]

Pipecolic and Piperidinyl PNAs: The pipecolyl PNA oligomers were designed by introducing a methylene, or an ethylene bridge between the C_γ- or C_β-carbon of the aminopropyl segment and the α''-carbon of the glycyl segment into an aminopropyl-glycyl PNA analogue (Figure 1.25).[105] The modified PNA-T_{10} oligomer (*pip*-PNA-**1**) carrying the *trans* (2*S*,4*S*) monomer destabilized the PNA$_2$:DNA triplex formation. However, the homothyminyl mixed *aeg*-PNA sequences incorporating the 5-aminopipecolyl unit (*pip*-PNA-**2**) form stable complexes with target DNA oligomers.
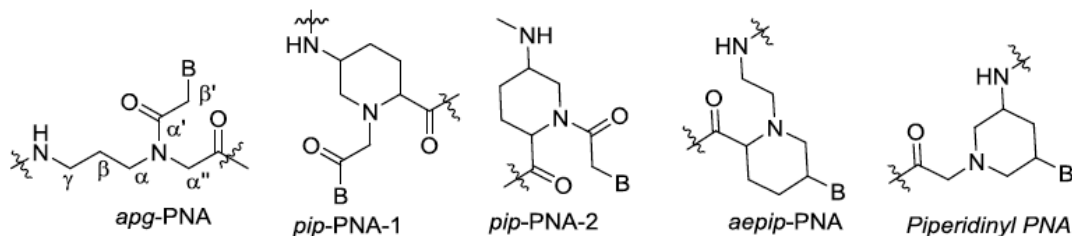
Figure 1.25 Pipecolic and Piperidinyl PNAs[102]

The chiral six-membered analogues (2*S*,5*R*) *aepip*-PNA and the *trans* (3*S*,5*S*)-piperidinyl PNA-T upon incorporation into the *aeg*-PNA-T_8 homooligomer at different positions stabilized the corresponding PNA$_2$:DNA triplexes.[106]

Cyclohexyl PNA: One of the earliest PNA modifications was to constrain the flexibility in the aminoethyl segment around a single C-C bond by introducing a cyclohexyl ring (Figure 1.26).[107] The derived *trans*-(*S*,*S*)-cyclohexyl PNA oligomer hybridized with complementary DNA as good as the unmodified *aeg*-PNA, while the enantiomeric *trans*-(*R*,*R*)-cyclohexyl PNA oligomers lacked such a property. Significant stereo differentiation was observed with *SR*- and *RS-ch*-PNAs, the *SR* isomer was more destabilizing than the *RS* isomer in *ch*- PNA:DNA complexes while In case of RNA complexes *RS* being more destabilizing than *SR*. From T_m values it was observed that (*R*,*S*)-*ch*-PNAs bound to RNA with higher affinity than to DNA. This is attributed to the inherent rigidity of the *cis-ch*-PNAs that forbids structural readjustments to bind to DNA

(PNA:DNA β ~ 140°) and prefers binding to RNA (PNA:RNA). This discrimination was achieved *via* the concept of preorganization.

Cyclopentyl PNAs: Ganesh *et al.*[108] also synthesized *cis*-(1*S*,2*R*/1*R*,2*S*)-cyclopentyl PNA-T oligomers in which the characteristic *endo-exo* puckering that dictates the pseudoaxial/pseudoequatorial dispositions of substituents may allow better torsional adjustments to attain the necessary hybridization-competent conformations. The homothymine *aeg*-PNA-T_{10} oligomers with modifications at defined positions exhibited binding affinities with DNA and RNA. It was observed that the *RS-cp*-PNA enantiomer showed higher affinity for DNA as compared to *SR-cp*-PNA isomer and in case of RNA this was reversed.
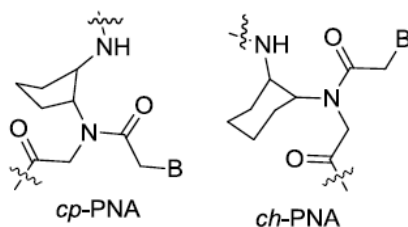
Figure 1.26 Cyclohexyl and cyclopentyl PNAs

The all-modified homooligomers of both enantiomers exhibited significant stabilization of their triplexes with DNA and poly rA without sacrificing the base specificity. *ch*-PNA and *cp*-PNA-T monomers of both (*S*,*R*) and (*R*,*S*) enantiomers were introduced into mixed sequences, in all the cases PNA:RNA hybrids were more stable than the corresponding PNA:DNA hybrids, and significantly, *cp*-PNA oligomers showed much higher T_ms compared to *ch*-PNA.

1.5.4 Synthetic heterocycles as *Janus* nucleobases

The term Janus heterocycle, was first coined by Nobel laureate J.M. Lehn, after the two-faced Roman god.[109,110] The term Janus heterocycle is employed to describe heterocycles that are capable of forming hydrogen bonds on both faces of the heterocycle,[110] in contrast to natural nucleobases that typically present only one face for functional hydrogen bonding (not counting the potential for simultaneous formation of Hoogsteen pairs) (Figure 1.27).
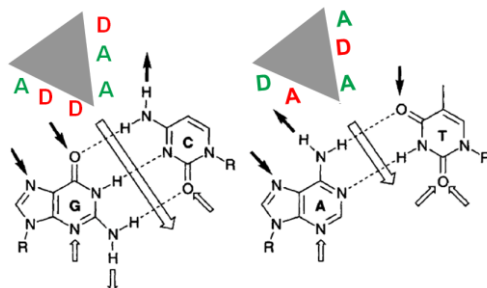
Figure 1.27 Janus-wedge recognizing regular DNA base pairs (A: hydrogen bond acceptors, D: hydrogen bond donors, R=deoxyribose)

A Janus heterocycle therefore represents one type of artificial molecule that can be rationally designed with predefined hydrogen bonding donors and acceptors relative to regular nucleobases. For example, if one side is complementary to A (which has the character of T) the other side can complement C (which has the character of G base), this Janus heterocycle is named Janus-TG. Typically, this Janus-TG heterocycle can form Watson-Crick base with A on one strand and C on another strand, as AC cannot form direct associate through base pairing. Taking this complementarity into account, there are two kinds of Janus heterocycles, one of which can form base pair with itself whereas the other cannot. So far, several Janus heterocycles have been reported, including the reported Janus-GA (Lehn *et al.*[111]) and subsequently, the Janus-GC (Mascal *et al.*[112]) and Janus-AT (Perrin *et al.*[113]) (Figure 1.28).
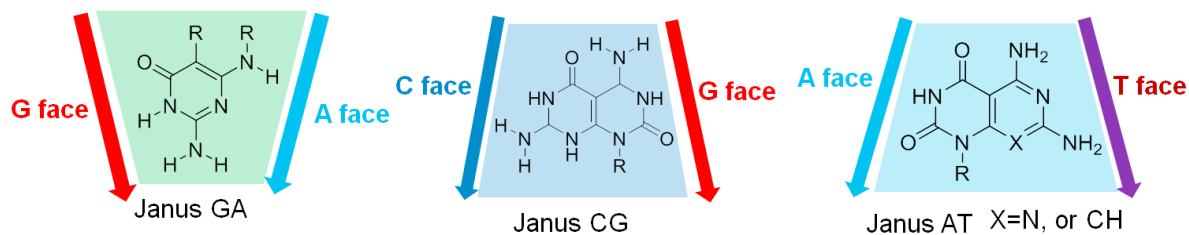
Figure 1.28 Janus heterocycles (R=alkyl groups in all three kinds of Janus heterocycles)

Due to self-complementary hydrogen-bonding donor and acceptor groups on two faces, Janus-GC and Janus-AT can aggregate to form interesting non-covalent supramolecules, the driving force being the formation of hydrogen bonding as well as stacking interactions that favor creation of higher order structures. For example, Fenniri *et al.*[114] reported the occurrence of helical rosette nanotubes in water by employing Janus-GC heterocycle (Figure 1.29b).
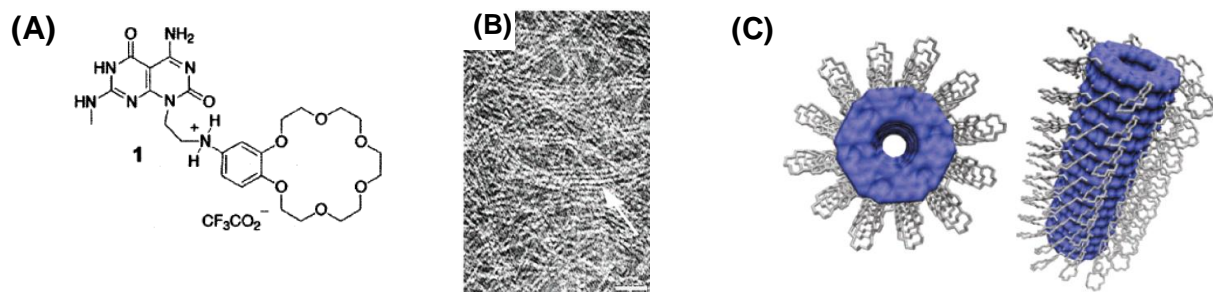
Figure 1.29 Self-assembly rosette nanotubes of Janus GC conjugate **(A)** GC conjugate 4-aminobenzo-18-crown-6-ether **(B)** TEM image of **A** **(C)** hierarchical self-assembly of **A**.

This group also reported that the property of helical rosette nanotubes could be adjusted by changing the side chain attached to the Janus-GC heterocycle. It was shown that chirality of rosette nanotubes could be induced with the addition of chiral amino acids when the side chain of Janus-GC was a crown ether. In addition, when Janus-GC heterocycle was combined with the ribose forming Janus-GC nucleosides, Janus-GC nucleosides also have the capacity to self-aggregate forming the aforementioned rosette nanotubes but with chiral properties owing to the presence of the ribose moiety.[115] When the Janus-AT nucleoside[115] was examined under similar conditions, instead of rosette cyclic structure, a linear sheet-like structure was observed. Although a given Janus-GC heterocycle or Janus-GC nucleoside can form the rosette nanotube superstructure, one significant limitation to the application of these kinds of rosette nanotubes is the lack of reproducibility of the identical nanotubes.[116] In order to achieve pre-defined nanotubes, they hypothesized that a certain number of Janus heterocycles should be incorporated into oligonucleotides to form a Janus heterocycle oligonucleotide, where sequence specificity can be modulated through suitable length as well as the incorporation of standard nucleosides. This is now being pursued by several laboratories.[117] Besides its utility in the construction of nanotubes, another potential application of a Janus heterocycle would be to act as a device by which one could recognize specific DNA sequences based on Watson-Crick base pairings.

1.5.4a Peptide Nucleic Acids: incorporation of Janus wedge motifs. PNA recognizes single-stranded and double-stranded nucleic acids through hydrogen bonding via Watson-Crick and Hoogsteen faces of the nucleobases. This approach works well for triplexes T-A-T, C$^+$-G-C, A-A-T, and G-G-C but does not cover the full scope of possibilities, such as mismatched nucleobases, due to difficulty in triplex formation with pyrimidine bases. McLaughlin *et al*[118] successfully used an oligo-PNA with Janus wedge motifs (Figure 1.30a) to target T-C

mismatches in DNA. Using a heterocycle called W which presents hydrogen bonding motifs that would complement the Watson-Crick faces of thymine and cytosine, they were able to form stable triplexes with a duplex DNA having 8 consecutive T-C mismatches flanked by eleven canonical bps. In later work (Figure 1.30a) heterocycles targeting A-T and (Figure 1.30b), G-C bps were able to successfully form triplexes when canonical bps were flanked by multiple mismatched bases and discriminate against non-complementary nucleobases.[119] Ganesh *et al*[120] utilized cyanuric acid (figure 13c) moiety displayed on a PNA backbone to strengthen triplex formation with oligo-A DNA. They found that a single cyanuric acid moiety in the middle of an oligo-T displaying PNA stabilized the triplex by a 12 °C in T_m measurements by providing a Watson-Crick competent binding face, without mattering which rotamer of the cyanuric acid presented to the DNA nucleobase.
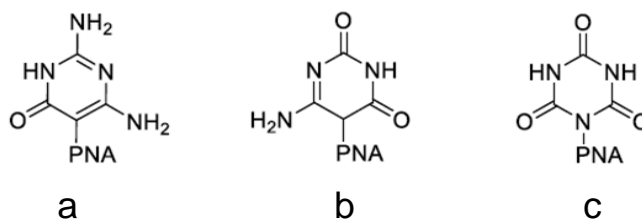
Figure 1.30 *Janus* wedge motifs used in PNA

1.5.4b Peptide Nucleic Acids: native peptide backbone. After Nielsen discovered PNA as a competent nucleic acid mimic, other groups have looked at native peptide backbones as a way of displaying genetic information. Lenzi *et al*[121] first attempted a native peptide backbone nucleic acid with a self-complementary sequence incorporating thymine and adenine nucleobases, but the hairpin had a modest T_m of 19 °C, possibly due to a lack of flexibility in the backbone. Diederichsen *et al*[122] synthesized alanyl peptide nucleic acids with alternating stereo centers to create nucleobase displaying beta sheets, again with modest success compared to PNA. Huang *et al*[123] introduced α-PNA in which the nucleobases are displayed on an alpha-helical peptide. Using a positively-charged helical structure, cooperative binding to the DNA sequence GGAGG was demonstrated as well as high melting temperatures in a possible triplex structure.

Eschenmoser and Krishnamurthy,[124] interested in the primordial evolution of genetic material, studied α-peptides and peptoids displaying triazine heterocycles as potential precursors to modern nucleic acids (Figure 1.31). In their systems 2,4-diamino-1,3,5-triazine and 2,4-dioxo-1,3,5-triazine rings were displayed on a glutamyl, aspartyl iminodiacetic acid, and

ethylenediamine monomers with alternating aspartic acid or dialkyl-ammonium groups for solubility. Their findings showed that the binding of these small oligos, displaying as few as six nucleobase mimics, were robust in binding to RNA and DNA when containing the 2,4-diamino-1,3,5-triazine heterocycle displayed on the longest linker. They found that the 2,4-dioxo-1,3,5-triazine nucleobase mimic did not bind to nucleic acid targets due to deprotonation in the assay buffer. While they ruled out that the triazine heterocycles as possible nucleobase precursors, they did show that α-peptides are competent as nucleic acid analogs even with negatively charged backbones. Furthermore, while not a major finding of the paper, they found that the 2,4-diamino-1,3,5-triazine is effective as a nucleobase mimic capable of recognizing uracil and thymine. In a follow up communication using α-peptides, the researchers showed that pK_a of the nucleobase is important.[125] Their findings yielded a relationship between pK_a of the base and pH of the solution, in that if the base becomes deprotonated or tautomerize in the buffer (at pK_a near or equal to the pH of the solution), then molecular recognition is weak.
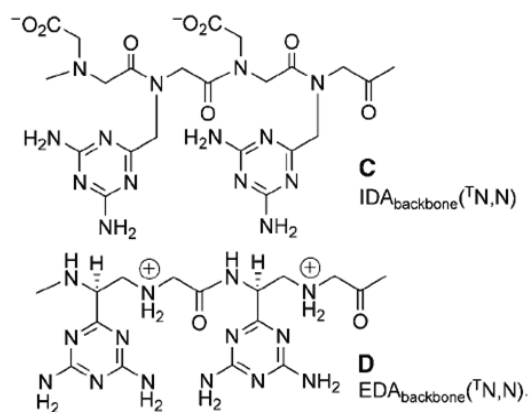
Figure 1.31 Structures of triazine-tagged oligomers (IDA=iminodiacetic acid; EDA=ethylenediamine; (^{T}N,N)=(2,4-diamino)triazin-6-yl)[124]

The purines (A and G) are capable of forming base pairing from both faces, through Watson-Crick / Hoogsteen H-bonding schemes, while pyrimidines (C and T) cannot achieve the same. In a way, A and G are "Janus" bases with two non-identical faces and are important in the formation of triplexes (A/T*A:T and C$^+$/G*G:C) and G-quadruplexes. Mutual recognition of non-complementary bases has been achieved by insertion of wedge-like heterocyclic "Janus bases" having two faces designed with H-bonding sites, matching that of the target nucleobases. The versatility of synthetic "Janus" bases to simultaneously recognize two natural nucleobases

and the geometric separation of adjacent bases in *aeg*-PNA attuned to that in natural DNA / RNA duplexes motivated us to design "Janus" PNAs.

1.5.5 Biological applications of PNAs

This section describes applications PNA and modified PNAs in gene therapeutics, molecular biology, and programmable supramolecular assemblies.

*1.5.5a **PNA as potential antisense and antigene agents**.* PNA can be used as an effective gene therapeutic agent by unique strand invasion property and their chemical as well as biological stability. There are basically two strategies involved in using PNAs as therapeutic drugs, namely antigene and antisense methods. Moreover, no sign of any general toxicity of PNA has so far been observed.

Inhibition of transcription. PNAs can inhibit the transcriptional processes by virtue of their ability to form a stable triplex, duplex of duplex structure or a strand invasion or strand displacement complex with DNA. Such complexes can create a structural hindrance to block the stable functioning of RNA polymerase and thus are capable of working as antigene agents.[123] Nielsen *et al*. have demonstrated that even an 8-mer PNA (T$_8$) is capable of blocking phage T$_3$ polymerase activity.[127]

Inhibition of translation In the case of antisense strategy, the nucleic acid analogs can be designed to recognize and hybridize to complementary sequences in *m*RNA and thereby inhibit its translation.[126] It has been established from the results of *in vitro* translation experiments involving rabbit reticulocyte lysates that both duplex-forming (mixed sequence) and triplex-forming (pyrimidine rich), PNAs can be useful for inhibiting translation at targets overlapping the AUG start codon.[128]

Inhibition of replication PNA can also inhibit the elongation of DNA primers by DNA polymerase. Further, the inhibition of DNA replication should be possible if the DNA duplex is subjected to strand invasion by PNA under physiological conditions or if the DNA is single-stranded during the replication process. Efficient inhibition of extra-chromosomal mitochondrial DNA, which is largely single-stranded during replication by PNA, has been demonstrated by Taylor *et al*.[129]

Inspite of these proof of concept experiments, till now no PNA is in clinical trials for any potential antisense applications.

1.5.6 Biophysical and biological properties of bifacial PNAs

Recently bifacial *Janus* bases (that bind one face by Watson-Crick and second face by non-natural H-bonding) have been used for invasion of double helical DNA in specific and stable manner by Ly et al.[130] Bivalent nucleic acid ligands have been recently used for recognition of RNA repeated expansion associated with Huntington's Disease by Ly et al.[131] In the recent work Bong et al.[132] describe a versatile non-covalent strategy based on triplex hybridization of oligo-uridylate RNA with bifacial polymer nucleic acid (bPoNA). bPoNA was prepared and side chain-functionalized with N-acetylgalactosamine (GalNAc), which is known to enable delivery to hepatocytes and liver via binding to the asialoglycoprotein receptor (ASGPR) cell.

2.0 Methodologies: Synthesis and Biophysical techniques

This section describes the method of solid phase synthesis of PNA oligomers, resin cleavage, purification, characterization and various methods for studying hybrid formation with cDNA such as temperature dependent absorbance, circular dichroism spectroscopy and isothermal titration calorimetry.

2.1 Solid phase synthesis of PNA oligomers

Peptides can be synthesized either in solution or by using solid phase synthesis protocols.[133] Synthesis of short peptides can be more efficient by solution phase strategy; however it requires tedious separation and purification steps after each coupling reaction. On the other hand, solid phase peptide synthesis can be efficiently used in the synthesis of several short and long chain peptides as well as in the synthesis of PNA oligomers.

Solid phase peptide synthesis, first invented by Merrifield,[134] utilizes polymeric beads, which has functional groups located on their surfaces and in their pores[135] (Figure 1.32). Small solid beads are insoluble in organic solvents and act as solid support. The peptide chains are grown over them by sequentially linking the amino acid monomers to the functional groups or linkers on the bead. In each step, the resin beads are immersed in the appropriate solvent containing the reagents for the reaction to proceed. Following the reaction, the solvent, excess

reagents and the byproducts are washed away by filtration. The next N^α-protected amino acid is coupled to the resin-bound amino acid either by using 3-hydroxy-2,3-dihydro-4-oxo-benzotriazole (DHBt) ester or by *in situ* activation with carbodiimide reagents. Because the *C*-terminal amino acid is linked to the insoluble solid support, it also acts as a protection for the carboxylic acid during the synthesis. The excess amino acid is washed out and the deprotection and coupling reactions are repeated until the desired peptide sequence is achieved. Finally, the resin-bound peptide and the side chain protecting groups are cleaved in a global deprotection step.
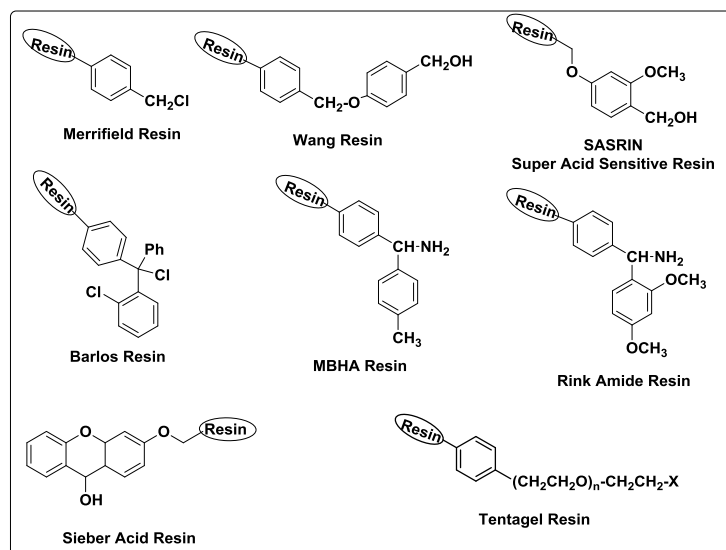
Figure 1.32 Representative structures of resin used in SPPS

There are two routinely followed protocols for solid phase peptide synthesis: *Fmoc* and *Boc* strategies, which use base labile and acid labile protecting groups respectively (Figure 1.33). The solid-phase peptide synthesis proceeds in C-terminal to N-terminal fashion. The N-terminus of amino acid monomers are protected by these two groups and added onto a deprotected amino acid chain. For N^α-protection, the first protocol uses the *t*-butoxy carbonyl (*t*-Boc) group, which can be removed by acidic conditions, such as 50% TFA in DCM. The reactive side chains are protected with groups that are stable to *t*-Boc deprotecting conditions and can be removed under strongly acidic conditions using HF[136] in dimethylsulfide or TFMSA[137] in TFA. In the alternative protocol, fluorenylmethyloxycarbonyl (Fmoc)[138] group is used for N^α-protection, it is stable to acidic conditions but can be cleaved off efficiently with a base such as piperidine.[139] The final peptide and side chain protecting groups can be cleaved with acid (50% TFA in DCM).
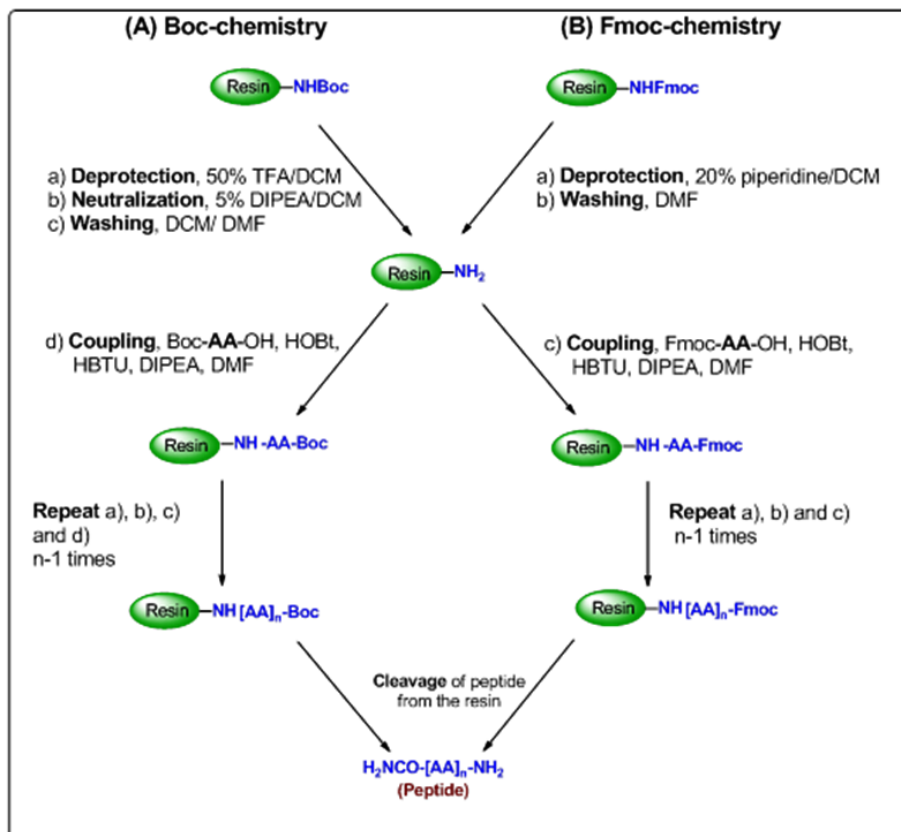
Figure 1.33 General protocols for SPPS via **(A)** Boc-chemistry **(B)** Fmoc-chemistry

The completely deprotected crude PNA oligomer released into solution is then checked by HPLC on reverse phase column and the desired product in the crude oligomer is identified by mass spectroscopy. It is then purified by semi-preparative HPLC and the purified PNA is checked by analytical HPLC and characterized by high resolution mass spectral measurements. These are the used for biophysical studies.

2.2 Biophysical techniques

The principles of various biophysical techniques used to study the properties of single-stranded *Janus* PNAs, *Janus* PNA:DNA triplex, *Janus* PNA:DNA duplex, double duplex of triplex, and *Janus* PNA:DNA double duplex complexes.

2.2.1 UV Job plot studies: Stoichiometry of binding

Job plot is a method to determine the stoichiometry of binding components in a binary complex.[140] In this continuous variation method,[141] the relative mole fractions of the two binding components are varied, keeping the total concentrations of two binding partners constant. The

UV absorption values of resultant component mixtures are plotted against the mole fractions of the two components where the maxima or minima on the plot correspond to the stoichiometry of the two binding species (Figure 1.34). There are several conditions that must be considered for Job method:

➢ The system must follow Lambert-Beer law within the absorbance range 0.1-1.0

➢ The total mole fractions of two components should be held constant throughout the experiment

➢ pH and ionic strength must remain constant
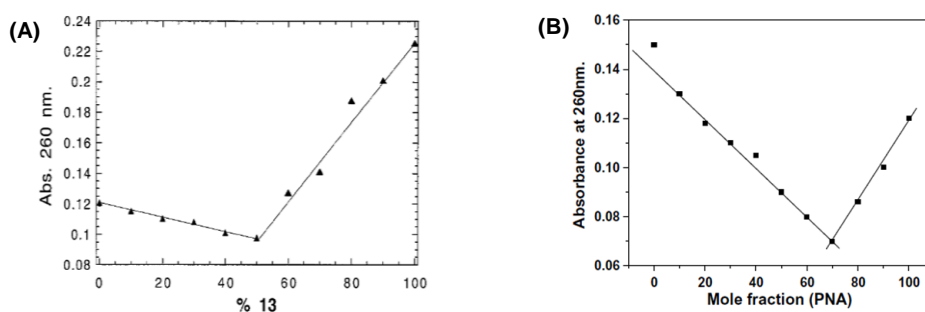
Figure 1.34 Job plot by UV continuation method of **(A)** duplex[142] **(B)** triplex[143]

In the present work, the binding of *Janus* PNA oligomers with one face of complementary DNA or simultaneously on both faces by respective complementary DNAs have been investigated by Job plot and temperature dependent UV spectroscopy for thermal stability.

2.2.2 Temperature dependent UV absorbance (UV-melting)

The two strands of complementary nucleic acids strands are held together by hydrogen bonding and the stacking interactions between the adjacent nucleobases. The stability of duplexes depends on temperature, pH and ionic strength. Under conditions that disrupt the hydrogen bonding and stacking interactions (eg. thermal, salt and pH), separation of complementary strands leads to denaturation of double helix. The separation of duplex into single strands is called melting and the thermal stability of various nucleic acid complexes including PNA-DNA/RNA hybrids can be monitored by increase in UV absorbance at 260 nm as a function of temperature.[144] The hetero-aromatic hydrophobic bases held together by H-bonding in duplex structure also interact via their π electron clouds leading to stacking in water. Because the UV absorbance of the bases is a consequence of π electron transitions, base stacking affects the magnitude of these transitions. The denaturation or dissociation of two strands results in loss

of stacking and disruption of hydrogen bonding which leads to the increase in absorbance at 260 nm (hyperchromicity).
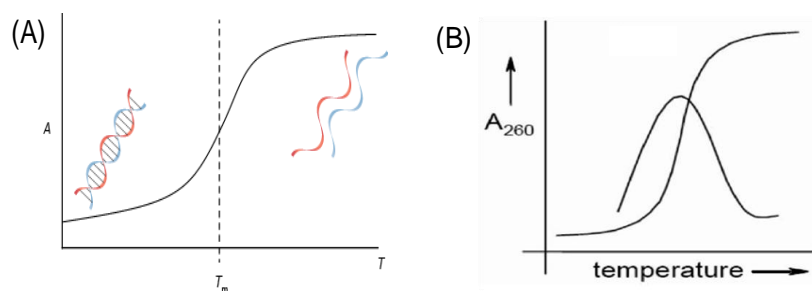

Figure 1.35 Temperature dependent UV absorbance[145] curves at 260 nm (**A**) sigmoidal graph and (**B**) derivative plot of duplex.

The process of DNA melting is co-operative and the plot of absorbance at 260 nm vs temperature is sigmoidal (Figure 1.35A). The sigmoidal nature of transition suggests that the nucleic acids exist in two states, either as duplex or as single strands and at varying temperatures, the relative proportions of two states change. A non-sigmoidal (e.g. sloping linear) transition with low hyperchromicity is suggestive of single strand/random conformation. The midpoint of the sigmoidal transition is termed as the melting temperature T_m. When the transitions are broad, the exact T_m values are obtained from the maximum of peak in the first derivative plots (Figure 1.35B). The T_m represents the thermal stability of duplexes, defined by the strength of base-dependent complementary interactions in nucleic acid hybrids. The binding stoichiometry of nucleic acids can be determined from the UV titration and the CD-mixing using Job plot. The combination of UV absorption and the CD spectra provides evidence for complex formation and the strand stoichiometry. In literature, it has been shown that polypyrimidine PNA binds to complementary DNA in a 2:1 ratio forming PNA_2:DNA triplex,[145] whereas the mixed purine-pyrimidine PNA binds in 1:1 ratio forming PNA:DNA duplex.[146] In this chapter PNAs are composed of both homo and mixed purine-pyrimidine bases and depending upon on the nature of sequence, *Janus* PNA and DNA form either triplex (2:1) or duplex (1:1) structures.

2.2.3 Circular Dichroism

Circular Dichroism (CD) is a spectroscopic technique used to measure the difference between the absorption of left-handed circularly polarized light (L-CPL) and the right-handed circularly polarized light (R-CPL) when they pass through a chiral medium.[146] The chiral molecules exist as isomers of non-superimposable mirror images called enantiomers. The

physical and chemical properties of enantiomeric pairs are identical except the ways by which they interact with circularly polarized light and other chiral molecules to form diastereomers. CD spectra are particularly valuable in determining the following aspects:

➢ Self-coiled or helical structures formation by individual strands

➢ Change of conformation upon binding to complementary oligonucleotide strand

➢ Conformations of hybrid PNA:DNA and DNA:DNA complexes

➢ Nature of helices: α, π (proteins) and β−sheet, A, B (DNA)

CD spectroscopy provides a reliable estimation of the overall conformational state of biopolymers and structural changes induced by modification as compared with the reference compound. In the case of nucleic acids, the nucleobases (chromophores) are attached to sugar residues that are chiral and hence show intense CD spectral signature of characteristic conformation (Figure 1.36A). CD spectroscopy monitors the structural changes of nucleic acids in solution and helps to diagnose whether new or unusual structures are formed by particular polynucleotide sequences. In the case of PNAs, the molecules are achiral and do not exhibit recognizable CD spectra. However, when they are chemically modified by substitutions on backbone or form hybrids with chiral DNA/RNA, they exhibit CD spectra (Figure 1.36B).
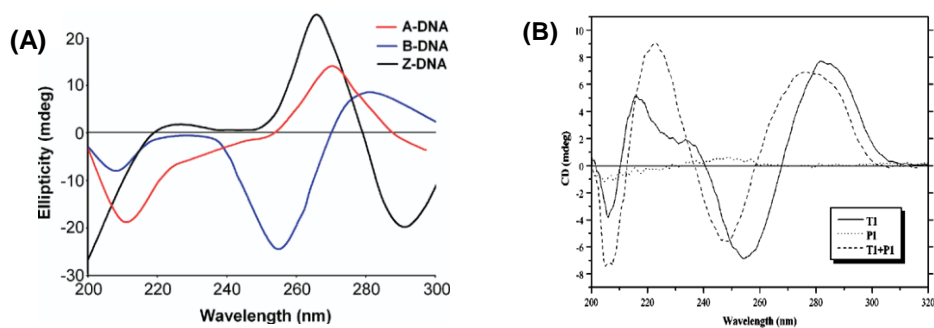
Figure 1.36 Typical CD spectra of **(A)** A, B and Z-DNA[147] **(B)** PNA:DNA hybrid [148]

2.2.4 Electrospray Ionization Mass Spectrometer (ESI MS) of Complex

More recently, mass spectrometry has emerged as an effective technique for identifying lead compounds on the basis of the characterization of noncovalent interaction. Electrospray (ESI) is an ionization technique to transform high molecular weight proteins and their complexes with ligands to gas phase for mass spectral analysis, It is ideal for characterizing intact protein-protein, protein-DNA, protein-RNA and DNA-DNA complexes and thereby defining

macromolecular and ligand-binding stoichiometry of complexes.[5] Biomolecules such as proteins that are noncovalently associated in solution with other ligands could be transferred into the gas phase with ESI and detected as an intact complex.[149]

In ESI-MS techniques, use of nitrogen gas jet in a stainless-steel capillary result in ion formation. From there, the ions are directed through focusing octupoles and lenses towards an ion trap device, where they are stored for subsequent mass measurement. *The linear or 3D-ion traps possess 4 parallel rod-shaped electrodes to separate the ions based on their molecular weights for their subsequent detection.* Quadrupole Time of Flight Mass Spectrometer (QTOF-MS)[150] is a kind of mass analyser which is a soft ionization technique capable of providing mass of both protonated and deprotonated molecules. These combine high sensitivity and mass accuracy for both precursor and product ions, in a rapid and efficient way.
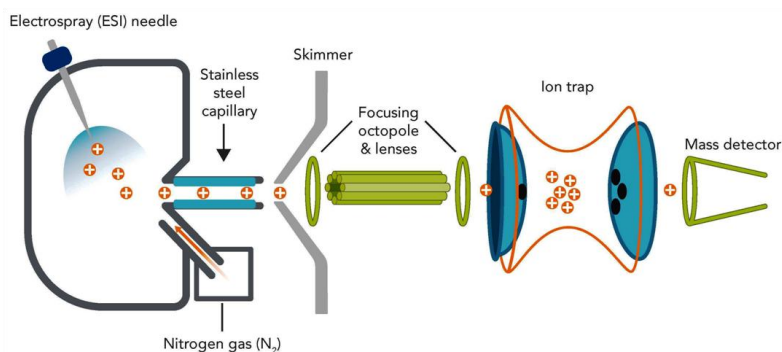
Figure 1.37 General component and mechanism of an ESI ion trap MS-instrument
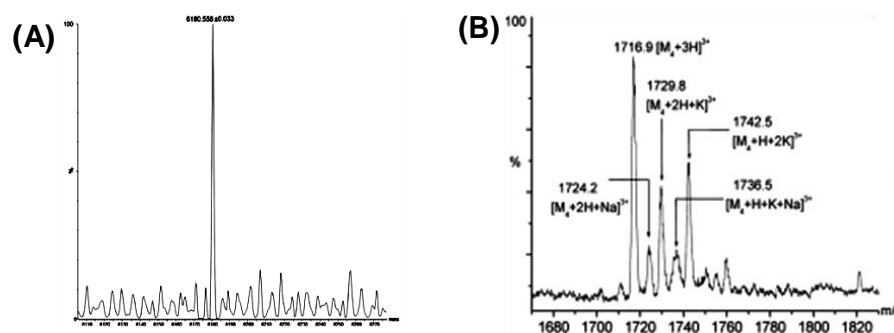
Figure 1.38 Example of noncovalently complex **(A)** ESI-MS spectrum of nucleopeptide /RNA duplex[151] **(B)** Partial nano-ESI-MS spectrum of PNA4 Quadruplex[152]

2.2.5 Isothermal Titration Calorimetry (ITC)

The strength of interactions between the two species in complex formation is represented by binding or association constant (K_a) or its inverse dissociation constant (K_d).[153] These can be

obtained by titrating one component to another and following the extent of complexation by different analytical techniques such as UV, CD, NMR etc. The thermodynamic details (enthalpy/entropy) of binding reaction is assessed by Isothermal Titration Calorimetry (ITC) technique in which the heat of a reaction produced during the bimolecular binding is measured at constant temperature during titration of one component into another.

In a typical ITC experiment, small aliquots of a titrant (eg. DNA, ligand etc)[154] solution is added to an analyte (eg. proteins, DNA, PNA etc)[155] solution at specific temperatures (below T_m duplex when the analyte PNA being titrated with complementary DNA) and the released heat is monitored. Judicious selection of nucleic acid concentrations and buffer conditions results in a titration curve that can be analyzed to yield the stoichiometry of the association reaction (N), the enthalpy of association (ΔH), the equilibrium association constant (K), and thus the free energy of association (ΔG). Once ΔH and ΔG are known, the entropy of association (ΔS) can also be obtained. Thus, a single ITC experiment yields a wealth of thermodynamic information about the association reaction, since there are unequal number of nucleobases, that binds in cooperative manner and showed 2 distinct transition (Figure 1.39).[156]
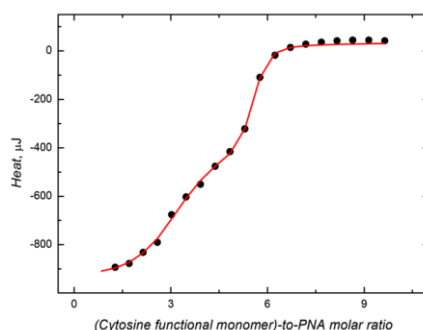
Figure 1.39 Typical ITC isotherm for 2-step binding complex[156]

3.0 Present work: Rationale

In the PNA literature so far, a single strand of PNA binds to a single strand of cDNA/RNA to form 1:1 duplexes or two strands of PNA bind to one strand of polypurine DNA strand to form triplexes. Both PNA and DNA or their other analogues carry one set of nucleobases on each strand to form duplexes involving one set of Watson-Crick base pairing (1:1 binding) (Figure 1.40A). Triplexes originate from three strands in which the central polypurine strand can form WC base pairing from one side and Hoogsteen base pairing from other side of purine(Figure 1.40B). In the present work, one strand of PNA is designed to have two sequence

of nucleobases on a single backbone, so that one PNA strand can form two duplexes, arising from two sequence of nucleobases attached to same strand. Such PNA termed "*Janus* PNA" has the potential to recognize two different DNAs, each DNA sequence complementing one or the other sequence. In these PNA oligomers, two nucleobases, are linked to one PNA monomer (Figure 1.40D).
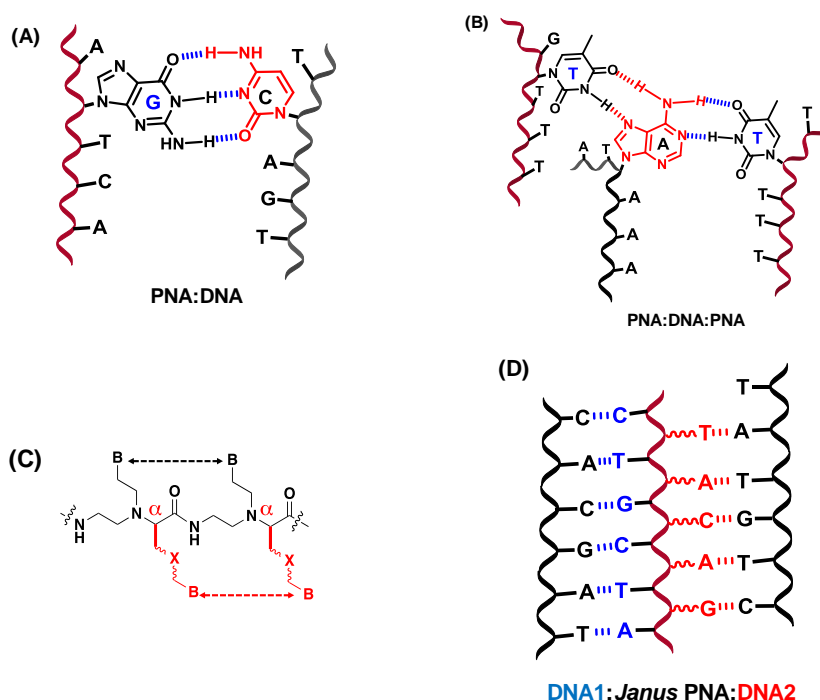

Figure 1.40 (A) PNA:DNA duplex **(B)** PNA$_2$:DNA triplex **(C)** *Janus* PNA oligomer having equi-internucleobases distance from both face **(D)** DNA **1**:*Janus* PNA:DNA **2** double duplexes.

3.1 *Janus* Peptide Nucleic Acid (*Janus* PNA)

We have designed *Janus* PNAs that contain an additional natural nucleobase (A, G, C and T) bearing side chain in each unit. In such PNAs, the backbone itself acquires "Janus" character and a single strand *Janus* PNA can now simultaneously recognize and bind two complementary DNA / RNA / PNA strands from both faces of the backbone by specific Watson Crick H-bonding. This would lead to formation of double duplexes through canonical Watson-Crick H-bonded base pairing on both faces from 3 strands. Herein we demonstrate the formation and characterization of double duplexes from single strand of *Janus* PNA that simultaneously binds to two complementary DNA strands (Figure 1.42).

The target *Janus* PNA structures consist of nucleobases (A/G/C/T) linked to C_α of glycyl component of standard *aeg*-PNA monomer in *S*-configuration, through a spacer link having triazole moiety. This arises from conjugation chemistry of precursor PNA having C_α-ethylazide side chain with N-propynyl nucleobases (A/G/C) using click chemistry. The *Janus* PNA thus has mixed sequence of base on one side linked via tertiary amide as in standard PNA and different base sequence on another side linked to C_α of glycyl via triazole. The synthetic strategy involved a combination of sequential coupling of C_α-ethylazido PNA-A/G/C/T monomers followed by click coupling and use of pre-clicked PNA monomers containing both the bases (C_α-ethyltriazolyl-A-*aeg* PNA-T). The tertiary amide side has 7 bases, while the triazolyl side has only 5 bases, for no particular reason other than to ensure good duplex formation and demonstrate the proof of concept (Figure 1.41c).

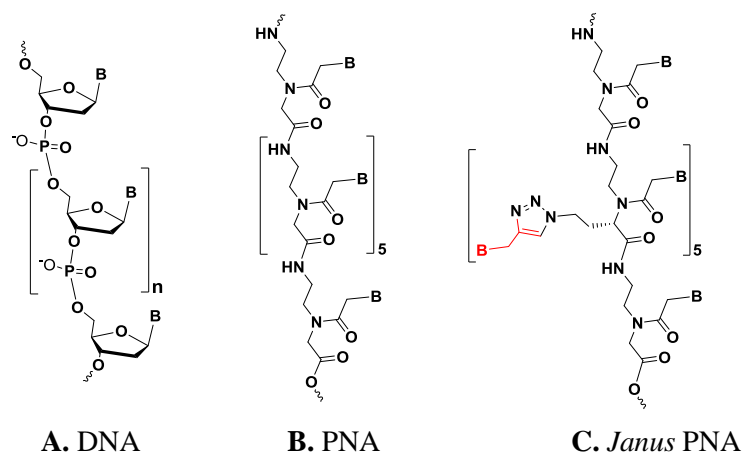**A.** DNA   **B.** PNA   **C.** *Janus* PNA

Figure 1.41 Structural comparison of DNA, PNA and *Janus* PNA (**B** = nucleobase A/T/G/C)
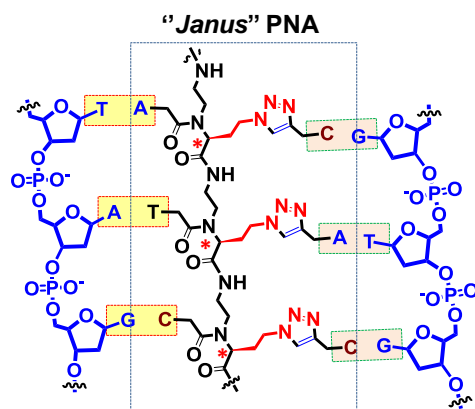

Figure 1.42 *Janus* PNA double duplexes (simultaneous binding with two different complementary DNA via Watson-Crick H-bonding)

Elucidation of the double helical structure of DNA, a genetic material and its remarkable properties has fascinated chemists. Tremendous efforts have been devoted to understanding its structure and function in biology as well as from a chemical and physicochemical standpoint. Investigations of oligonucleotides as potential therapeutics that target nucleic acids has led the search for nucleic acid mimetics with improved properties. However, several generic problems remain to be solved towards the medicinal applications of such functional nucleic acids and their derivatives, although there are already 5 drugs approved by FDA. These include degradation by cellular nucleases, impermeability through cell membrane, low hybridization efficiency as well as targeting two different genes simultaneously in specific and stable manner, by utilization of Watson-Crick pairing between the heterocyclic bases.
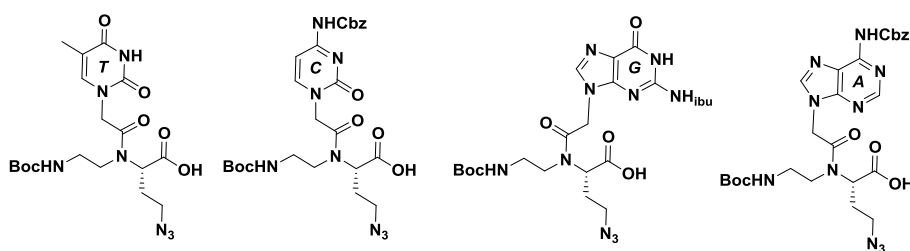
3.2 Scope of present work

The preceding section has given an overview on the structure directed function of nucleic acids and their analogues especially on peptide nucleic acid (PNA) which is an explored, but promising DNA mimic. PNA has properties of natural oligonucleotides and therefore has great potential for the development of therapeutics in the form of antisense/antigene oligonucleotides. The major drawbacks of PNA like poor water solubility, inefficient cellular uptake, self-aggregation and ambiguity in directional selectivity of binding restrict its applications in terms of designing PNA-based gene-targeted drugs and showing week supramolecular self-assembly. Hence, various modifications of PNA to overcome these limitations have been employed by different research groups as described in the previous section. Bifacial nucleobases, that are synthetic can also perform similar functions as PNAs and can be used for targeting two mismatches in gene and stop its function. This adds another dimension to PNA applications.

The work presented in this thesis deals with an entirely new and novel design of dual presentation of natural nucleobases (not synthetic bifacial heterocycles) on either side of PNA backbone to make PNAs that can simultaneously recognize two different strands of cDNA/RNA/PNA by standard Watson-Crick H-bonding. The synthesis and characterization of such *Janus* PNA with natural nucleobases anchored on either side a single backbone and offering two faces for target recognition is presented in this thesis. The hybridization properties of such PNAs to form two WC duplexes from same PNA backbone, their conformation by CD and

energetics of their binding is investigated. Further exploration of i-motif and G-tetraplexes from such *Janus* PNAs and morphological features of their supramolecular assembly are studied.

3.2.1 Chapter 2: Syntheses and characterization of C_α-substituted *aeg*-PNA monomers as precursors for the assembly of *Janus* PNAs
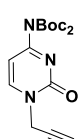
This chapter aims to address the design, synthesis of C_α-ethylazido substituted *aeg* PNA monomers (**4**-**8**) as rationally designed PNA analogs having side chains carrying azido function for conjugating different nucleobases (eg. **9**-**10**) by click reaction in solution. The propargyl nucleobases (**11**-**13**) can also be conjugated post assembly of PNA sequences by click reaction on solid phase synthesis. The presence of nucleobases on both faces synergistically increase the *Janus* PNA-DNA duplex stability of each face, and presence of both faces complementary DNA complex stabilized.

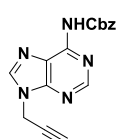**4.** α-(*S-eaz*) *aeg* T   **5.** α-(*S-eaz*) *aeg* C   **6.** α-(*S-eaz*) *aeg* G   **7.** α-(*S-eaz*) *aeg* A PNA
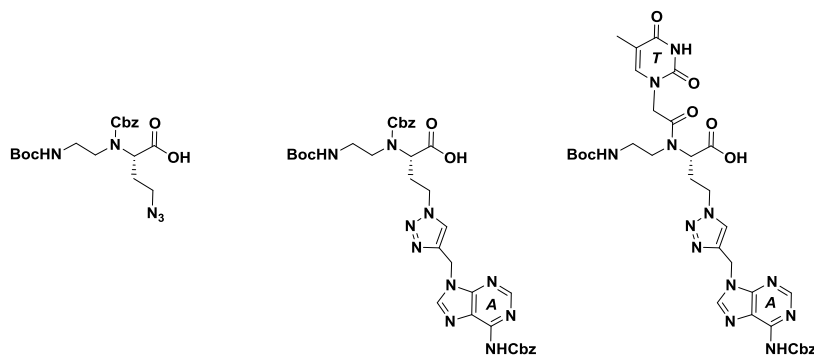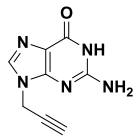
8. α-(*S*-eaz)]-N^{Cbz}-*aeg*   **9.** α-(*S*-etz -N9-A^{NHCbz})-N^{Cbz} *aeg*   **10.** α-(*S*-etz-N9- A^{NHCbz})-*aeg*-T

11   **12**   **13**

3.2.2 Chapter 3: Solid Phase Synthesis, Purification and Characterization of *Janus* PNA Oligomers

Chapter 3 deals with incorporation of C_α-modified PNA units by solid phase followed by global click reaction to synthesize the various kinds of *Janus* Peptide nucleic acids oligomers, - *homo Janus* PNA, *chimeric Janus* PNA, *Hetero Janus* PNA or *Self complementary mix Janus* PNAs, depending on types of bases and linkage connecting them to backbone – ex amide or triazole. The synthesis of PNA oligomers were followed by cleavage from resin, purification by RP-HPLC and characterization by MALDI-TOF spectrometry.
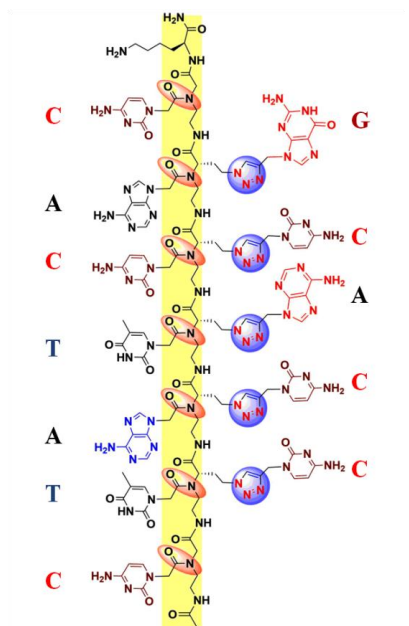
Hetero Janus PNA oligomer (**JP 6**)

3.2.3 Chapter 4: Biophysical evaluation of modified PNA oligomers

This chapter reports on the biophysical evaluation of various kinds of *Janus* PNA Oligomers. UV Job plot by continuous variation method was used to determine stoichiometry. Temperature-dependent UV absorbance studies lead to determination of thermal stabilities of *Janus* PNA:DNA duplexes, triplexes, double duplexes and duplex of triplex with complementary antiparallel DNA. The effect of various substitutions at C_α - position on the conformation of *Janus* PNA:DNA duplexes, triplexes, double duplexes and duplex of triplex is investigated by Circular Dichroism (CD) studies. The thermodynamics studies of *Janus* PNA:DNA duplexes, triplexes, double duplexes and duplex of triplex is investigated by Isothermal Titration

Calorimetry (ITC) experiments. And the noncovalent interaction in the gas phase of *Janus* PNA:DNA duplexes, triplexes, double duplexes and duplex of triplex studies with ESI-MS is also described.

3.2.4 Chapter 5: i-Motif and G-quadruplex as well as supramolecular assembly studies of *Janus* PNA Oligomers

Chapter 5 deals with studies on the pK_a determination of N3 of C (cytidine) in PNA C-oligomers as it is crucial in formation of C-C$^+$ i-motif by *homo Janus* PNA and its stability studies by pH-dependent UV-T_m, comparison with i-motifs formed by DNA, *aeg*-PNA, amino-triazolyl PNA. Studies on G quadruplex formation of *Janu*s PNA and its thermal stability are compared with tetraplexes from corresponding DNA, *aeg*-PNA and amino-triazolyl PNAs at different salt concentrations.

The morphological studies of *Janus* PNAs along with its complex formation with cDNA, individually with amide face, triazole face and together studied by FESEM technique are described.

4.0 References

1. Alberts, B.; Johnson, A.; Lewis, J.; Raff, M.; Roberts, K.; Wlater, P. *Molecular Biology of the Cell* (4th Ed.). *Garland Science.* **2002**, pp. 120-121.

2. Watson, J. D.; Crick, F. H. C. *Nature*. **1953**, *171*, 737-738.

3. Srinivasan, A. R.; Sauers, R. R.; Fenley, M. O.; Boschitsch, A. H.; Matsumoto, A.; Colasanti, A. V. Olson, W. K. *Biophys. Rev.* **2009**, *1*, 13.

4. Hoogsteen, K. *Acta. Crystal.* **1963**, *16*, 907-916.

5. (a) Crick, F. H. C. *J. Mol. Biol*. **1966**, *19*, 548-555. (b) Soll, D.; Cherayil, J. D.; Bock, R. M.; *J. Mol. Biol.* **1967**, *29*, 97-112.

6. (a) Dickerson, R. E. *Methods in Enzymol*. **1992**, *211*, 67-111. (d) Saenger, W. Principles of Nucleic Acids structure. Springer-Verlag, New York, **1984**. (b) Wang, A. H. J.; Quigley, G. J.; Kolpak, F. J.; Van der M. G.; Van Boom, J. H.; Rich, *Science*, **1981**, *211*, 171-176.

7.(a) Strobel, S. A.; Doucette-Stamm, L.A.; Riba, L.; Housman, D.E.; Dervan, P.B.; *Science*, *254*, 1639-1642. (b) Huesken, D.; Deichert, A.; Hall, J.; Haner, R. *Nucleosides & Nucleotides*

1999, *18*, 1507-1511. (c) Haner, R.; Hall, J.; *Antisense Nucleic Acid Drug Dev.* **1997**, *7*, 423-430.

8. (a) Toule, J-J. in Antisense oligonucleotides and antisense RNA: Novel Pharmacological and Therapeutic Agents. (Weiss, B., ed), CRC press, **1997**, 1-16. (b) Wagner, R. W. *Nat. Med.* **1995**, *1*, 1116-1118.

9. Chan, P. P.; Glazer, P. M. *J. Mol. Med.* **1997**, *75*, 267-282.

10. Stephenson, M. L.; Zamecnik, P.C. *Proc. Natl. Acad. Sci. U.S.A* **1978**, *75*, 285-288.

11. Uhlmann, E.; Peyman, A. *Chem. Rev.* **1990**, *90*, 4, 543-584.

12. Bennett, C. F.; Swayze, E. E. *Annu. Rev. Pharmacol. Toxicol.* **2010**, *50*, 259-293.

13. Stein, C. A.; Cohen, J. S. Phoshorothioate oligodeoxynucleotide analogues. In Cohen J. S. (ed.): Oligodeoxynucleotides-Antisense Inhibitors of Gene Expression. London: Macmillan Press, **1989**, p. 97.

14. Millar, P. S. Non-ionic antisense oligonucleotides. In Cohen, J. S. (ed.): Oligodeoxynucleotides-antisense inhibitors of gene expression. London: Macmillan Press, **1989**, p. 79.

15. Froehler, B.; Ng, P.; Matteucci, M. *Nucleic Acids Res.* **1988**, *16*, 4831-4839.

16. Summers, M. F.; Powell, C.; Egan, W.; Byrd, R. A.; Wilson, W. D.; Zon, G. *Nucleic Acids Res.* **1986**, *14*, 7421-7437.

17. (a) Sood, S.; Shaw, B. R.; Spielvogel, B. F. *J. Am. Chem. Soc.* **1990**, *112*, 9000-9001. (b) Shaw, B. R.; Madison, J.; sood, S.; Spielvogel, B. F; In Agrawal, S. (ed): *Methods in Molecular biology*, vol 20: Protocols for Oligonucleotides and Analogs. Synthesis and properties. Totowa, NJ. Humana Press, Inc. **1993**, 225-243. (c) Sergueev, D. S.; Shaw, B. R. *J. Am. Chem. Soc.* **1998**, *120*, 9417-9427.

18. Uhlmann, E.; Peyman, A. *Chem. Rev.* **1990**, *90*, 4, 543-584.

19. Brown, D. A.; Kang, S. H.; Gryaznov, S. M.; De Dionisio, L.; Heidenreich, O.; Sullivan, S.; Xu, X.; Neerenberg, M. I. *J. Biol. Chem.* **1994**, *43*, 26801-26805.

20. Guvakova, M. A.; Yakubov, L. A.; Vlodavsky, I.; Tonkinson, J. L.; Stein, C. A. *J. Biol. Chem.* **1995**, *270*, 2620-2627.

21. Rockwell, P., O'Connor, W.; King, K.; Goldstein, N. I.; Zhang, L. M.; Stein, C. A. *Proc. Natl. Acad. Sci. USA,* **1998**, *94*, 6523-6528.

22. Bennett, C. F.; Swayze, E. E. *Annu. Rev. Pharmacol. Toxicol.* **2010**, *50*, 259-293.

23. Altona, C.; Sundaralingam, M.; *J. Am. Chem. Soc.* **1972**, *94*, 8205-8212.

24. Teplova, M.; Minasov, G.; Tereshka, V.; Inamati, G. B.; Cook, P. D. *et al. Nat. Struct. Biol.* **1999**, *6*, 535-539.

25. (a) Braasch, D. A.; Corey, D. R. *Chem. Biol.* **2001**, *8*, 1-7. (b) Orum, H.; Wengel, J. *Curr. Opin. Mol. Ther.* **2001**, *3*, 239-243.

26. Bondensgaard, K.; Petersen, M.; Singh, S. K.; Rajwanshi, V. K.; Kumar, R.; Wengel, J.; Jacobsen, J. P. *Chem. Eur. J.* **2000**, *6*, 2687-2695.

27. Braasch, D. A.; Liu, Y.; Corey, D. R. *Nucleic Acids Res.* **2002**, *30*, 5160-5167.

28. Seth, P. P.; Siwkowski, A.; Allerson, C. R.; Vasquez, G.; Lee, S. *et al. J. Med. Chem.* **2009**, *52*, 10-13.

29. Koizumi, M. *Curr. Opin. Mol. Ther.* **2006**, *8*, 144-149.

30. Sazani, P.; Gemignani, F.; Kang, S. H.; Maier, M. A.; Manoharan, M. et al. *Nat. Biotechnol.* **2002**, *20*, 1228-1233.

31. Alter, J.; Lou, F.; Rabinowitz, A.; Yin, H.; Rosenfeld, J. *et al. Nat. Med.* **2006**, *12*, 175-177.

32. Wang, J.; Verbeure, B.; Luyten, I.; Lescrinier, E.; Froeyen, M.; Hendrix, C.; Rosemeyer, H.; Seela, F.; van Aerschot, A.; Herdewijn, P. *J. Am. Chem. Soc.* **2000**, *122*, 8595-8602.

33. Verbeure, B.; Lescrinier, E.; Wang, J.; Herdewijn, P. *Nucleic Acids Res.* **2001**, *29,* 4941-4947.

34. Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science*, **1991**, *254*, 1497-1500.

35. Egholm, M.; Buchardt, O.; Nielsen, P. E.; Berg, R. H. *J. Am. Chem. Soc.* **1992**, *114*, 1895-1897.

36. Egholm, M.; Nielsen, P. E.; Buchardt, O.; Berg, R. H. *J. Am. Chem. Soc.* **1992**, *114*, 9677-9678.

37. Brown, S. C.; Thomson, S. A.; Veal, J. M.; Davis, D. G. *Science*, **1994**, *265*, 777-780.

38. Rasmussen, H.; Kastrup, J. S.; Nielsen, J. N.; Nielsen, J. M.; Nielsen, P. E. *Nat. Struct. Biol.* **1997**, *4*, 98-101.

39. (a) Leijon, M.; Graeslund, A.; Nielsen, P. E.; Buchardt, O.; Norden, B.; Kristensen, S. M.; Eriksson, M. *Biochemistry*, **1994**, *22*, 9820-9825. (b) Eriksson, M.; Nielsen, P. E. *Nat. Struct. Biol.* **1996**, *3*, 410-413.

40. Bets, L.; Josey, J. A.; Veal, J. M.; Jordan, S. R. *Science*, **1995**, *270*, 1838-1841.

41. (a) Eriksson, M.; Nielsen, P. E.; *Q. Rev. Biophys.* **1996**, *29*, 369-394. (b) Nielsen, P. E.; Egholm, M. *Curr. Iss. Molec. Biol.* **1999**, *1*, 89-104.

42. Uhlmann, E.; Will, D. W.; Breipohl, G.; Langner, D.; Ryte, A. *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 2632-2635.

43. Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science*, **1991**, *254*, 1497-1501.

44. (a) Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science*, **1991**, *254*, 1497-1501. (b) Nielsen, P. E.; Egholm, M.; Buchardt, O. *J. Mol. Recogn.* **1994**, *7*, 165-170.

45. (a) Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Anti-Cancer Drug Design.* **1993**, *8*, 53-63. (b) J. C. Hanvey, N. J. Peffer, J. E. Bisi, S. A. Thomson, R. Cadilla, J. A. Josey, D. J. Ricca, C. F. Hassman, M. A. Bonham, K. G, Au, S. G. Carter, D. A. Bruckenstein, A. L. Boyd, S. A. Noble, L. E. Babiss, *Science*, **1992**, *258*, 1481-1485.

46. Nielsen, P. E. *Acc. Chem. Res.* **1999**, *32*, 624-630.

47. Davis, J. T. *Angew. Chem., Int. Ed.* **2004**, *43*, 668-698.

48. Simonsson, T. *Biol. Chem.* **2001**, *382*, 621-628.

49. Egholm, M.; Buchardt, O.; Christensen, L.; Behrens, C.; Freier, S. M.; Driver, D. A.; Berg, R. H.; Kim, S. K.; Norden, B.; Nielsen, P. E. *Nature,* **1993**, *265*, 566-568.

50. Krishnan-Ghosh, Y.; Stephens, E.; Balasubramanian, S. *J. Am. Chem. Soc.* **2004**, *126*, 5944-5945.

51. Datta, B.; Schmitt, C.; Armitage, B. A. *J. Am. Chem. Soc.* **2003**, *125*, 4111-4118. (b) Marin, V. L.; Armitage, B. A. *J. Am. Chem. Soc.* **2005**, *127*, 8032-8033.

52. Sharma, N. K.; Ganesh, K. N. *Chem. Commun.* **2005**, 4330-4332.

53. Demidov, V. V.; Frank-Kamenetskii, M. D.; *Trends Biochem. Sci.* **2004**, *29,* 62-71.

54. Nielsen, P. E, *Pure Appl. Chem.* **1998**, *70*, 105-110.

55. Hyrup, B.; Nielsen, P. E. *Bioorg. Med. Chem.* **1996**, *4*, 5-23.

56. Koppelhus, U.; Nielsen, P. E. *Adv. Drug Delivery Rev.* **2003**, *55*, 267-280.

57. Kumar, V. A.; Ganesh, K. N. *Acc. Chem. Res.* **2005**, *38*, 404-412.

58. Corradini, R.; Sforza, S.; Tedeschi, T.; Totsingan, F.; Manicardi, A.; Marchelli, R. *Curr. Top. Med. Chem.* **2011**, *11*, 1535-1554.

59. Dueholm, K. L.; Petersen, K. H.; Jensen, D. K.; Egholm, M.; Nielsen, P. E.; Buchardt, O. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 1077-1080.

60. Sforza, S.; Corradini, R.; Ghirardi, S.; Dossena, A.; Marchelli, R. Eur. *J. Org. Chem.* **2000**, *2000*, 2905-2913.

61. Menchise, V.; de Simone, G.; Tedeschi, T.; Corradini, R.; Sforza, S.; Marchelli, R.; Capasso, D.; Saviano, M.; Pedone, C. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 12021-12026.

62. Koppelhus, U.; Awasthi, S. K.; Zachar, V.; Holst, H. U.; Ebbesen, P.; Nielsen, P. E. *Antisense Nucleic Acid Drug Dev*. **2002**, *12*, 51-63.

63. Zhou, P.; Dragulescu-Andrasi, A.; Bhattacharya, B.; O'Keefe, H.; Vatta, P.; Hyldig- Nielsen, J. J.; Ly, D. H. *Bioorg. Med. Chem. Lett*. **2006**, *16*, 4931-4935.

64. Haaima, G.; Lohse, A.; Buchardt, O.; Nielsen, P. E. *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 1939-1942.

65. Puschl, A.; Sforza, S.; Haaima, G.; Dahl, O.; Nielsen, P. E. *Tetrahedron Lett.* **1998**, *39*, 4707-4710.

66. Aguado, G. P.; Rua, F.; Branchadell, V.; Nielsen, P. E.; Ortuno, R. M. *Tetrahedron: Asymmetry* **2006**, *17*, 2499-2503.

67. Gourishankar, A.; Ganesh, K. N. *Artificial DNA PNA XNA* **2012**, *3*, 5-13.

68. Sugiyama, T.; Imamura, Y.; Demizu, Y.; Kurihara, M.; Takano, M.; Kittaka, A. *Bioorg. Med. Chem. Lett*. **2011**, *21*, 7317-7320.

69. Kosynkina, L.; Wang, W.; Liang, T. C. *Tetrahedron Lett*. **1994**, *35*, 5173-5176.

70. Tedeschi, T.; Sforza, S.; Corradini, R.; Marchelli, R. Tetrah*edron Lett.* **2005**, *46*, 8395-8399.

71. Englund, E. A.; Appella, D. H. *Org. Lett.* **2005**, *7*, 3465-3467.

72. Dose, C.; Seitz, O. *Org. Lett.* **2005**, *7*, 4365-4368.

73. Ficht, S.; Dose, C.; Seitz, O. *ChemBioChem.* **2005**, *6*, 2098-2103.

74. Englund, E. A.; Wang, D.; Fujigaki, H.; Sakai, H.; Micklitsch, C. M.; Ghirlando, R.; Martin-Manso, G. *Nat. Commun.* **2012**, *3*, 614.

75. Sforza, S.; Tedeschi, T.; Corradini, R.; Marchelli, R. *Eur. J. Org. Chem.* **2007**, *2007,* 5879-5885.

76. Avitabile, C.; Moggio, L.; Malgieri, G.; Capasso, D.; Gaetano, S. D.; Saviano, M.; Pedone, C.; Romanelli, A. *PLoS One* **2012**, *7*, e35774.

77. Dragulescu-Andrasi, A.; Rapireddy, S.; Frezza, B. M.; Gayathri, C.; Gil, R. R.; Ly, D. H. *J. Am. Chem. Soc*. **2006**, *128*, 10258-10267.

78. Yeh, J. I.; Boris Shivachev, B.; Rapireddy, S.; Crawford, M. J.; Gil, R. R.; Du, S.; Madrid, M.; Ly, D. H. *J. Am. Chem. Soc.* **2010**, *132*, 10717-10727.

79. Crawford, M. J.; Rapireddy, S.; Bahal, R.; Sacui, I.; Ly, D. H. J. *Nucleic Acids* **2011**, 2011, doi:10.4061/2011/652702.

80. Manicardi, A.; Fabbri, E.; Tedeschi, T.; Sforza, S.; Nicoletta Bianchi, N.; Brognara, E.; Gambari, R.; Marchelli, R.; Corradini, R. *Chem. Bio. Chem.* **2012**, *13*, 1327-1337.

81. Sahu, B.; Sacui, I.; Rapireddy, S.; Zanotti, K. J.; Bahal, R.; Armitage, B. A.; Ly, D. H. *J. Org. Chem.* **2011**, *76*, 5614-5627

82. Canady, T. D.; Telmer, C. A.; Oyaghire, S. N.; Armitage, B. A.; Bruchez, M. P. *J. Am. Chem. Soc.* **2015**, *137*, 10268-10275.

83. Mitra, R.; Ganesh, K. N. *J. Org. Chem.* **2012**, *77*, 5696-5704.

84. Shiraishi, T.; Pankratova, S.; Nielsen, P. E. *Chem. Biol*. **2005**, *12*, 923-929

85. Jain, D. R. and Ganesh, K. N. *J. Org. Chem*. **2014**, *79*, 6708-6714.

86. Tomac, S.; Sarkar, M.; Ratilainen, T.; Wittung, P.; Nielsen, P. E.; Norden, B.; Graslund, A. *J. Am. Chem. Soc*. **1996**, *118*, 5544-5549.

87. Hollenstein, M.; Leumann, C. *J. Org. Lett.* **2003**, *5*, 1987-1990.

88. Rose, D. J. *Anal. Chem.* **1993**, *65*, 3545-3549.

89. Bennett, C. F.; Swayze, E. E. *Annu. Rev. Pharmacol. Toxicol*. **2010,** 50, 259-293

90. Gangamani, B. P.; Kumar, V. A.; Ganesh, K. N. *Tetrahedron* **1996**, *52*, 15017-15030.

91. Gangamani, B. P.; D' Costa, M.; Kumar, V. A.; Ganesh, K. N. *Nucleosides Nucleotides* **1999**, *18*, 1409-1011.

92. Gangamani, B. P.; Kumar, V. A.; Ganesh, K. N. *Tetrahedron* **1999**, *55*, 177-192.

93. D'Costa, M.; Kumar, V. A.; Ganesh, K. N. *Org. Lett*. **1999**, *1*, 1513-1516.

94. Vilaivan, T.; Khongdeesameor, C.; Harnyuttanokam, P.; Westwell, M. S.; Lowe, G. *Bioorg. Med. Chem. Lett*. **2000,** *10*, 2541-2545.

95. D'Costa, M.; Kumar, V. A.; Ganesh, K. N. *Org. Lett.* **2001,** *3*, 1281-1284.

96. Sharma, N.; Ganesh, K. N. *Tetrahedron Lett.* **2004,** *45*, 1403-1406.

97. Sharma, N.; Ganesh, K. N. *Chem. Commun.* **2003**, *0*, 2484-2485.

98. Englund, E. A.; Xu, Q.; Witschi, M. A.; Appella, D. H. *J. Am. Chem. Soc.* **2006**, *128*, 16456 – 16457.

99. Puschl, A.; Boesen, T.; Zuccarello, G.; Dahl, O.; Pitsch, S.; Nielsen, P. E. *J. Org. Chem.*

2001, *66,* 707-712.

100. Hickman, D. T.; King, P. M.; Cooper, M. A.; Slater, J. M.; Mickelfield, J. *Chem. Commun*. **2000**, 2251-2252.

101. Kumar, V. A.; Pallan, P. S.; Meena,; Ganesh, K. N. *Org. Lett.* **2001**, *3*, 1269-1272.

102. Kumar, V. A.; Meena. *Nucleos Nucleot Nucleic Acids* **2003,** *22*, 1101-1104.

103. D'Costa, M.; Kumar, V. A.; Ganesh, K. N. *Tetrahedron. Lett*. **2002,** *43,* 883-886.

104. Lonkar, P.; Ganesh, K. N.; Kumar, V. A. *Org. Biomol. Chem*. **2004,** *2,* 2604-2611

105. Lonkar, P. S.; Kumar, V. A.; Ganesh, K. N. *Nucleos Nucleot Nucleic Acids* **2001**, *20*, 1197-1200.

106. Lescrinier, E.; Froeyen, M.; Herdewijn, P. *Nucleic Acids Res*. **2003**, *31*, 2975-2989.

107. Lagrioule, P.; Wittung, P.; Eriksson, M.; Jensen, K. K.; Norden, B.; Buchardt, O.; Nielsen, P. E. *Chem. Eur. J.* **1997**, *3*, 912-919.

108. Govindaraju, T.; Kumar, V. A.; Ganesh, K. N. *J. Org. Chem.* **2004**, *69*, 5725-5734.

109. Branda, N.; Kurz, G.; Lehn, J.-M. *Chem. Commun.* **1996,** *2*, 2443-2444.

110. Marsh, A.; Nolen, E. G.; Gardinier, K. M.; Lehn, J.-M. *Tetrahedron Lett.* **1994,** *35*, 397-400.

111. Fenniri, H.; Mathivanan, P.; Vidale, K. L.; Sherman, D. M.; Hallenga, K.; Wood, K. V.; Stowell, J. G. *J. Am.Chem. Soc.* **2001,** *123*, 3854-3855.

112. (a) Mascal, M.; Hext, N. M.; Warmuth, R.; Arnall-Culliford, J. R.; Moore, M. H.; Turkenburg, J. P. *J. Org. Chem.* **1999,** *64*, 8479-8484. (b) Mascal, Mark ; Hext, Nicholas M. ; Warmuth, Ralf ; Moore, Madeleine H. ; Turkenburg, Johan P. *Angew. Chem. Int. Ed.* **1996,** *35*, 2204-2206

113. Asadi, A.; Patrick, B. O.; Perrin, D. M. *J. Org. Chem.* **2007,** *72*, 466- 475.

114. (a) Moralez, J. G.; Raez, J.; Yamazaki, T.; Motkuri, R. K.; Kovalenko, A.; Fenniri, H. *J. Am. Chem. Soc.* **2005,** *127*, 8307-8309. (b) Fenniri, H.; Deng, B.-L.; Ribbe, A. E. *J. Am. Chem. Soc.* **2002,** *124*, 11064-11072. (c) Fenniri, H.; Deng, B.-L.; Ribbe, A. E.; Hallenga, K.; Jacob, J.; Thiyagarajan, P. *Proc. Natl. Acad. Sci. U.S.A* **2002,** *99*, 6487-6492.

115. (a) Zhao, H.; Huang, W.; Wu, X.; Xing, Z.; He, Y.; Chen, Q. *Chem Commun.* **2012,** *48*, 6097-6099. (b) Yang, H.-Z.; Pan, M.-Y.; Jiang, D.-W.; He, Y *Org. Biomol. Chem.* **2011,** *9*, 1516-1522. (c) Pan, M.-Y.; Wu, X.-H.; Luo, D.-B.; Huang, W.; He, Y. *Acta Crystallographica Section C* **2011,** *67*, o175-0178. (d) Pan, M.-Y.; Hang, W.; Zhao, X.-J.;

Zhao, H.; Deng, P.-C.; Xing, Z.-H.; Qing, Y.; He, Y. *Org. Biomol. Chem.* **2011,** *9*, 5692-5702.

116. Asadi, A.; Patrick, B. O.; Perrin, D. M. *J. Org. Chem.* **2006,** *72*, 466-475.

117. Beingessner, R. L.; Diaz, J. A.; Hemraz, U. D.; Fenniri, H. *Tetrahedron Lett.* **2011,** *52*, 661-664.

118. Chen, D.; Meena; Sharma, S. K.; McLaughlin, L. W. *J. Am. Chem. Soc.* **2004,** *126*, 70-71.

119. Chen, H.; Meena; McLaughlin, L. W. *J. Am. Chem. Soc.* **2008,** *130*, 13190-13191.

120. Vysabhattar, R.; Ganesh, K. N. *Tetrahedron Lett.* **2008,** *49*, 1314-1318.

121. Lenzi, A.; Reginato, G.; Taddei, M.; Trifilieff, E. *Tetrahedron Lett.* **1995,** *36*, 1717-1718.

122. Diederichsen, U. *Angew. Chem. Int. Ed. Engl.* **1996,** *35*, 445-448.

123. Huang, Y.; Dey, S.; Zhang, X.; Sonnichsen, F.; Garner, P. *J. Am. Chem. Soc.* **2004,** *126*, 4626-4640.

124. Mittapalli, G. K.; Reddy, K. R.; Xiong, H.; Munoz, O.; Han, B.; De Riccardis, F.; Krishnamurthy, R.; Eschenmoser, A. *Angew. Chem. Int. Ed.* **2007,** *46*, 2470-2477.

125. Mittapalli, G. K.; Osornio, Y. M.; Guerrero, M. A.; Reddy, K. R.; Krishnamurthy, R.; Eschenmoser, A. *Angew. Chem. Int. Ed.* **2007,** *46*, 2478-2484.

126. Ray, A.; Norden, B. *FASEB J.* **2000**, *14*, 1041-1060.

127. Nielsen, P. E.; Egholm, M.; Buchardt, O. *Gene* **1994**, *149*, 139-145.

128. Knudsen, H.; Nielsen, P. E. *Nucleic Acids Res.***1996** *24*, 494-500.

129. Taylor, R. W.; Chinnery, P. F.; Turnbull, D. M.; Lightowlers, R. N. *Nature Genet.* **1997**, *15*, 212-215.

130. Thadke, S. A.; Hridya, V. M.; Perera, J. D. R.; Gil, R. R.; Mukherjee, A.; Ly, D. H. *Commun. Chem.* **2018**, *1*, 79.

131. Hsieh, W.-C.; Bahal, R.; Thadke, S. A.; Bhatt, K.; Sobczak, K.; Thornton, C.; Ly, D. H. *Biochemistry* **2018**, *57*, 907-911.

132. Xia,X.; Zhou, Z.; DeSantis, C. ; Rossi, J. J.; Bong, D. *ACS Chem. Biol.* **2019**, *14*, 1310-1318.

133. (a) Amblard, M.; Fehrentz, J.-A.; Martinez, J.; Subra, G. *Mol. Biotechnol.* **2006**, *33*, 239-254. (b) Bodansky, M.; Bodansky, A. *Practice of Peptide Synthesis*, Springer-Verlag, Berlin, **1984**. (C) Stewart, J. M.; Young, J. D. *Solid Phase Peptide Synthesis*, W. H. Freeman & Co, New York, **1969**.

134.(a) Merrifield, R. B. *J. Am. Chem. Soc.***1963**, *85*, 2149–2154. (b) L. Connah, R. Joshi, S. Vibhute, G. Gambino, J. D. G. Correia, G. Angelovski, *Org. Lett.* **2019**, *21*, 5378-5382. (c) Agouridas, V.; El Mahdi, O.; Diemer, V.; Cargoet, M.; Monbaliu, J.-C. M.; Melnyk, O. *Chem. Rev.* **2019**, *119*, 7328-7443. (d) Isidro-Llobet, A.; Kenworthy, M. N.; Mukherjee, S.; Kopach, M. E.; Wegner, K.; Gallou, F.; Smith, A. G.; Roschangar, F. *J. Org. Chem.* **2019**, *84*, 4615-4628. (e) Kulkarni, S. S.; Wang, C.-C.; Sabbavarapu, N. M.; Podilapu, A. R.; Liao, P.-H.; Hung, S.-C. *Chem. Rev.* **2018**, *118*, 8025-8104.

135. Christensen, L.; Fitzpatrick, R.; Gildea, B.; Petersen, K.; Hansen, H. F.; Koch, C.; Egholm, M.; Buchardt, O.; Nielsen, P. E.; Coull, J.; Berg, R. H. *J. Peptide Sci.* **1995**, *3*, 175–183.

136. Muttenthaler, M.; Albericio, F.; Dawson, P. E. *Nat. Protoc.* **2015**, *10*, 1067.

137. Orain, D.; Ellard, J.; Bradley, M. *J. Comb. Chem.* **2002**, *4*, 1-16.

138. Hansen, P. R.; Oddo, A. in *Peptide Antibodies: Methods and Protocols* (Ed.: G. Houen), Springer New York, New York, NY, **2015**, pp. 33-50.

139. Brieke, C.; Cryle, M. J. *Org. Lett.* **2014**, *16*, 2454-2457.

140. (a) Barawkar D. A.; Bruice T. C. *Proc. Natl. Acad. Sci. U.S.A* **1998**, *95*, 11047-11052. (b) Yadav D. K.; Lokhande R. S.; Pitale S. M.; Janwadkar S. P.; Navarkar P. S.; Rana P. K. *J. Anal. Chem.* **2014**, *2*, 10-14.

141. (a) Huang, C. Y. *Methods in Enzymology, (*Ed.: D. L. Purich), Academic Press, **1982**, *87*, 509-525. (b) Nielsen, P. E.; Christensen, L. *J. Am. Chem. Soc.* **1996**, *118*, 2287-2288.

142. Barawkar, D. A.; Bruice, T. C. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 11047-11052.

143. Vysabhattar, R.; Ganesh, K. N. *Tetrahedron Lett.* **2008**, *49*, 1314-1318.

144. Puglisi, J. D.; Tinoco, I. Jr. *Methods Enzymol*. **1989**, *180*, 304-325.

145. Egholm, M.; Buchardt, 0.; Nielsen, P. E.; Berg, R. H. *J. Am. Chem. Soc.* **1992**, *114*, 1895-1897.

146. Egholm, M.; Buchardt, 0.; Christensen, L.; Behrens, C.; Frier, S. M.; Driver, D. A.;Berg, R. H.; Kim, S. K.; Norden, B.; Neilsen, P. E. *Nature*, 1993, *365*, 566-568.

147. Baker, E. S.; Bowers, M. T. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 1188-1195.

148. Kushon, S. A.; Jordan, J. P.; Seifert, J. L.; Nielsen, H.; Nielsen, P. E.; Armitage, B. A. *J. Am. Chem. Soc.* **2001**, *123*, 10805-10813.

149. (a) Kempen E. C.; Brodbelt J. S. *Anal. Chem.* **2000**, *72*, 5411-5416. (b) Ganem, B.; Li, Y. T.; Henion, J. D. *J. Am. Chem. Soc.* **1991**, *113*, 6294-6296.

150. Chernushevich, Igor V.; Loboda, Alexandre V.; Thomson, Bruce A. *Eur. J. Mass Spectrum.* **2001**, *36*, 849-865.

151. Mercurio, M. E.; Tomassi, S.; Gaglione, M.; Russo, R.; Chambery, A.; Lama, S.; Stiuso, P. Cosconati, S. Novellino, E.; Di Maro, S.; Messere, A. *J. Org. Chem.* **2016**, *81*, 11612-11625.

152. Krishnan-Ghosh, Y.; Stephens, E.; Balasubramanian, S. *J. Am. Chem. Soc.* **2004**, *126*, 5944-5945.

153. Gourishankar, A.; Shukla, S.; Ganesh, K. N.; Sastry, M. *J. Am. Chem. Soc.* **2004**, *126*, 13186-13187.

154. Pierce, M. M.; Raman, C. S.; Nall, B. T. *Methods* **1999**, *19*, 213.

155. (a) Barbieri, C. M.; Li, T. K.; Guo, S.; Wang, G.; Shallop, A. J.; Pan, W.; Yang, G.; Gaffney, B. L.; Jones, R. A.; Pilch, D. S. *J. Am. Chem. Soc.* **2003**, *125*, 6469. (b) Kunne, A.; Sieber, M.; Meierhans, D.; Allemann, R. K. *Biochemistry* **1998**, *37*, 4217. (c) Egholm, M.; Buchardt, 0.; Christensen, L.; Behrens, C.; Frier, S. M.; Driver, D. A.;Berg, R. H.; Kim, S. K.; Norden, B.; Neilsen, P. E. *Nature*, 1993, *365*, 566-568.

156. Bartold, K.; Pietrzyk-Le, A.; Golebiewska, K.; Lisowski, W.; Cauteruccio, S.; Licandro, E.; D Souza, F.; Kutner, W. *ACS Appl. Mater. Interfaces* **2018**, *10*, 27562-27569.

Chapter 2

Synthesis and Characterization of C_α Substituted *aeg*PNA Monomer Precursors for Assembly of Designed *Janus* PNAs

2.0 Introduction

As described in Chapter 1, functionalization of the PNA backbone may dramatically change the physicochemical properties of PNA and influence the bio-distribution affecting the pharmacokinetic profile. A number of efforts have been made to synthesize conformationally constrained cyclic PNA analogs using the concept of structural preorganization to address interesting attributes of PNA:DNA/RNA hybridization.[1] Further, modifications in the acyclic PNA backbone to make them chiral with stereocentres at C_α, or C_γ carbon[2] have led to improvement in the properties of derived PNAs (Figure 2.1). Introduction of gem-dimethyl substitution at C_α have led to PNA analogs that also show preferential binding to cDNA rather than cRNA observed earlier with cyclohexyl PNAs.[3] Recently, Sugiyama *et al.*[4] have introduced methyl substitution at the least studied C_β position. The functionalization at these positions of the monomer may also preorganize the PNA oligomer. Specifically, configuration of the new stereogenic centres (C_α, C_β and C_γ) has been reported to influence the preference of the PNA to form left-handed or right handed helices. (Figure 2.1). This, in turn, affects the stability of the PNA:DNA duplex through the control of helix handedness.[5]

I Achiral PNA   **II** Chiral α-PNA   **III** Chiral β-PNA   **IV** Chiral γ-PNA
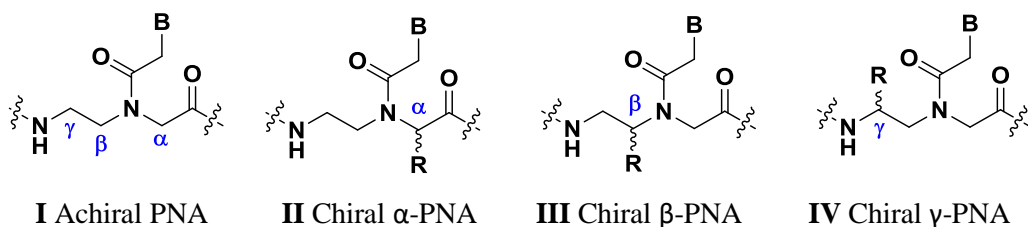
Figure 2.1 Structures of achiral PNA, chiralshow α-, β- and γ-modified PNA (R = amino acid side chain or other modifications, B = nucleobase)

In order to be a good DNA binding molecule, a modified PNA backbone should be neither too flexible nor too rigid. A highly flexible backbone would require a very large loss of entropy for efficient binding whereas a very rigid backbone would prevent DNA binding due to a difficult fit to adopt to target structure.[5] The ability of unmodified aminoethyl glycyl (*aeg*) PNA to bind strongly with DNA has been interpreted as a result of the 'constrained flexibility' in its structure[6] imposed by the tertiary amide link. The earliest and simplest modification involved in extension of PNA backbone by linking with a methylene group individually in each of the structural subunits (aminoethyl[7] and glycine[8,9]) of the PNA monomer. However, such modifications resulted in a significant decrease in the melting temperatures of the derived

PNA:DNA duplexes. These studies suggest that the inter-residue distance between the nucleobases is important and compatible with that present in DNA and RNA and is vital for efficient binding with complementary oligonucleotides.

Several PNA derivatives were obtained by insertion of side chains at either C_α- or C_γ-carbons on *aeg* PNA backbone. These include hydroxymethyl (serine), aminobutyl (lysine), guanidininopropyl (arginine) and many other functional groups from amino acid side chains.[9] Such acyclic chiral PNA analogues, retained the constrained flexibility of original PNA backbone, without much preorganization observed earlier with cyclic PNA analogues. However, the stereochemistry of the side chain substitutents at C_α and C_γ seems to be important for the derived PNA oligomers in exhibiting efficient hybridization properties: *R* configuration at C_α and *S* configuration at C_γ.[10] The side chains of the D/L amino acids are versatile for easy synthesis of such chiral substituted PNA oligomers. This chapter describes synthesis and characterization of various C_α-ethylazido substituted PNA monomers (Figure 2.3) that are precursors for assembling *Janus* PNA oligomers (Figure 2.2) either through the coupling of prior synthesized *Janus* triazolyl monomers for mixed sequence *Janus* oligomers or by post-assembly click reaction for *Janus* homo-oligomers.
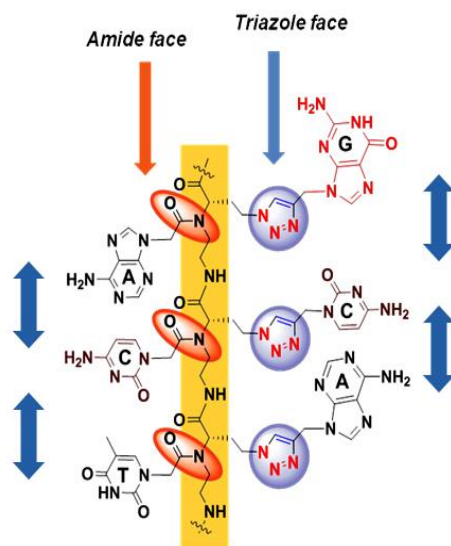
Figure 2.2 Representation of a *Janus* oligomers having an amide and a triazole face

1. α(*S*)-(eaz) *aeg* A/T/CG

2. α(*S*)-[(etz–N9(A^{NHCbz})]-N^{Cbz} *aeg*

3. α(*S*)-[(etz-N9(A^{NHCbz})]-N(acetamido thymnyl *aeg*
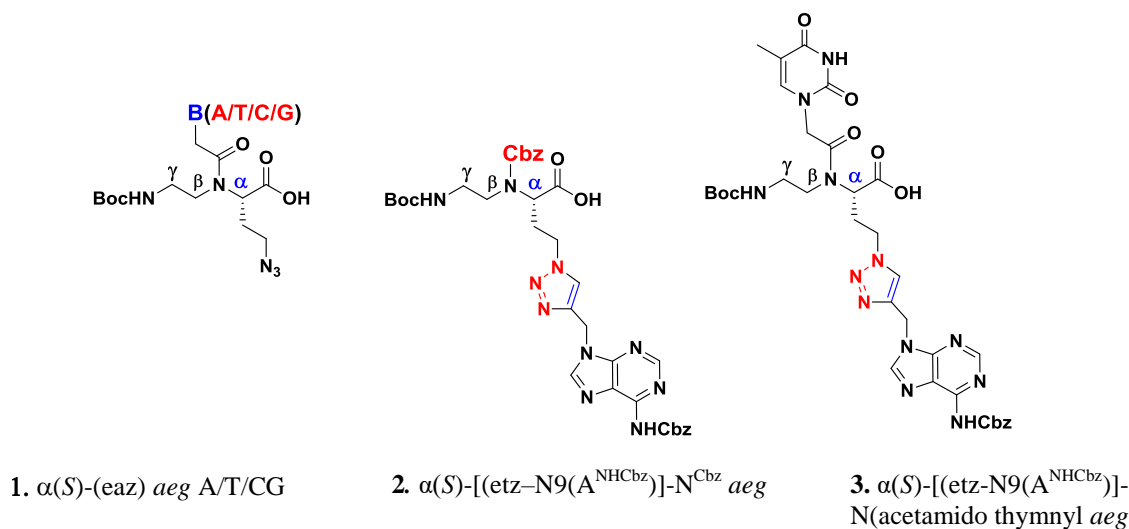
Figure 2.3 Target PNA and *Janus* PNA monomers modified at α-position

2.1 Rationale for the present study

The precursors for *Janus* PNA oligomers are *aeg* PNA (A/G/C/T) monomers (Figure 2.3) having ethylazido side chain at C$_\alpha$ that can be conjugated with A/G/C/T base propynes via click reaction.[11] The resulting triazolyl linker is less prone for nonspecific secondary interactions with nucleotide phosphate groups in the cDNA sequence. Such precursors also impart chirality to achiral PNAs and may have some stereospecific effects in complementation. The C$_\alpha$-ethylazido *aeg* (A/C/G/T) PNA monomers (**1**) can be incorporated in PNA sequences on solid phase followed by single global click reaction with any propargyl nucleobase to yield corresponding *Janus* PNAs that are homo-oligomeric on both faces or have mixed sequence on the amide face but homo-oligomer on the triazole face (Figure 2.4). Such kinds of *Janus* PNA oligomers possess one face of nucleobases attached by amide bond through a carbon linker as in standard PNA, linker (triazolyl face). Alternatively, the monomer precursor of *Janus* PNA having nucleobases at both amide and triazolyl faces can be synthesized by solution phase click reaction of C$_\alpha$-ethylazido *aeg* PNA (A/C/G/T) monomers with propargyl nucleobases and straightaway assembled to yield final *Janus* PNA oligomers. The latter strategy is versatile since it leads to *Janus* PNAs with mixed nucleobase sequence on both the faces (Figure 2.5). In these *Janus* PNAs, the inter-nucleobase geometric distance in neighboring units on either amide or triazolyl face is always the same corresponding to that in DNA and RNA. This facilitates simultaneous

complementary binding with DNA or RNA from both faces, such modified, *Janus* PNA also becomes chiral.
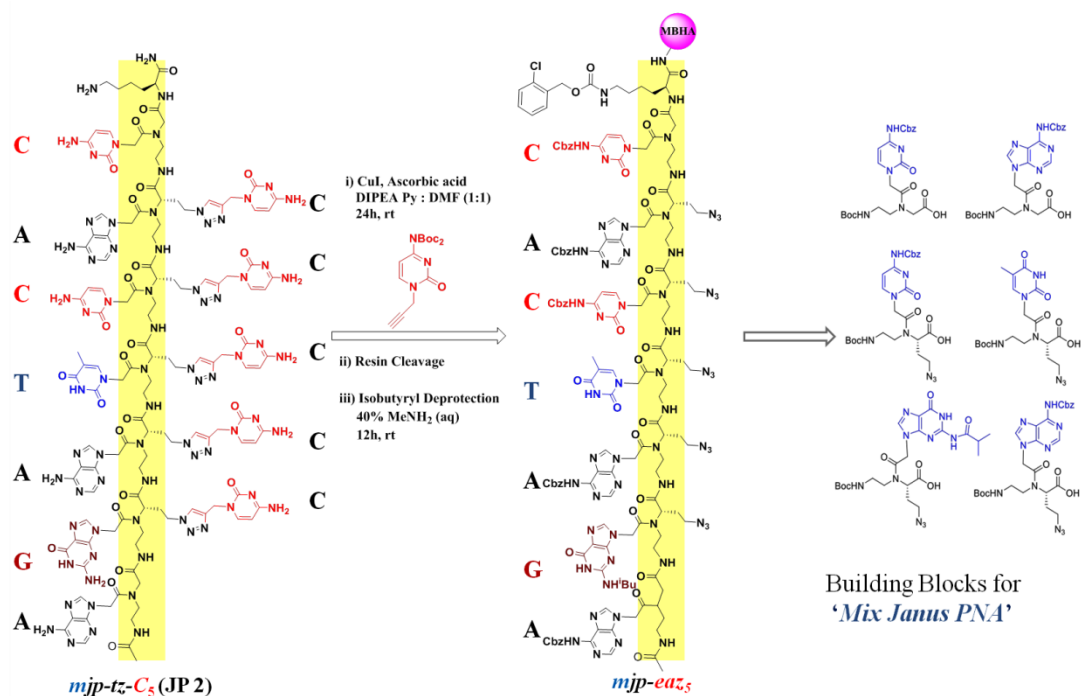
Figure 2.4 Retro synthetic strategy of **JP 2** by global click reaction
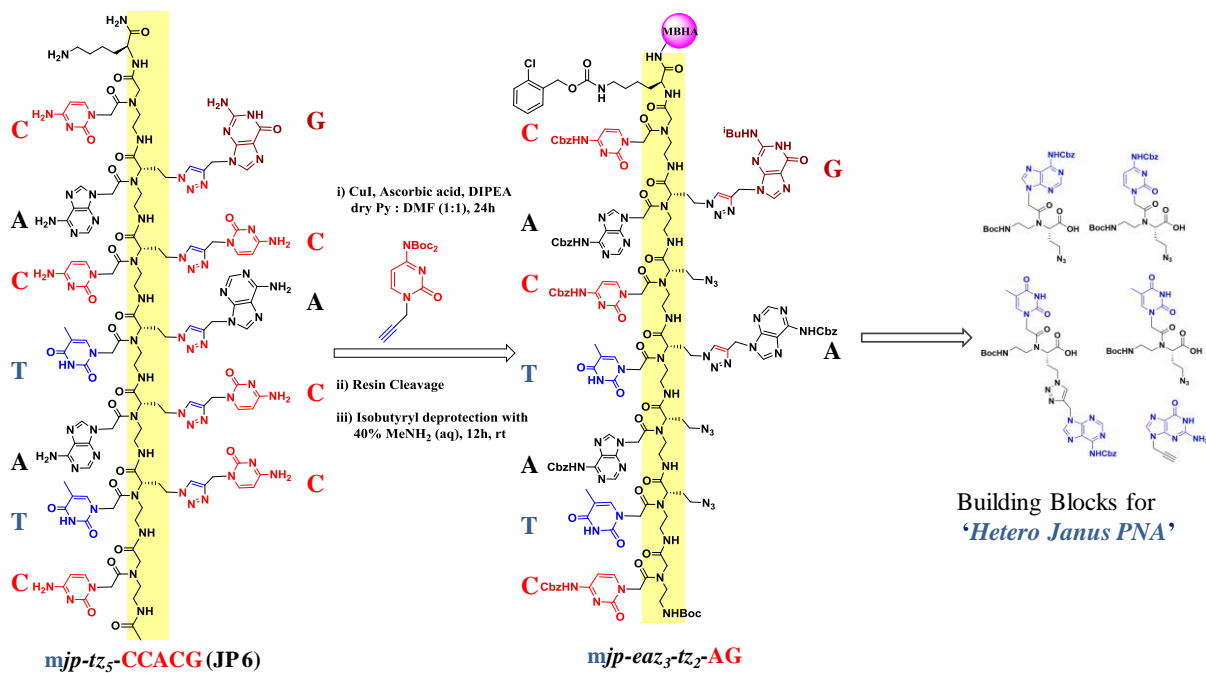
Figure 2.5 Retro synthetic strategy of **JP 6** from preclicked monomers

58

As far as the regio and stereospecificity of the chiral monomer is concerned, PNAs having monomers derived from D-amino acids (C_α-*R*) in the glycyl fragment of *aeg* backbone bind to the complementary DNA with greater thermal stability than those derived from the corresponding L-amino acids (C_α-*S*) in the glycyl fragment of *aeg* backbone. Substituted *eda* fragment in *aeg* backbone derived from L-amino acids translate into (C_γ-*S*) PNA oligomers and are known to preorganize into a right-handed helix, binding to complementary DNA/RNA with high affinity.[12] The situation turns out to be the reverse for substitutions of chiral amino acids in glycine segment in *aeg* backbone; the D-amino acids with C_α-*R* at glycine hybridize with cDNA better than that with L-amino acids (C_α-*S*) at glycine site.[13]

These results inspired the design and synthesis of a C_α-substituted PNA oligomers. These incorporate *S*-ethylazido substitution at C_α in the *aeg* PNA backbone to result α-(*S*-ethylazido) aminoethylglycyl A/T/C/G PNA monomers (**4-7**) required for solid phase synthesis of *Janus* PNA homooligomers by post solid phase single click reaction. The monomers (**4**) can also be coupled with propargyl nucleobases (**11**) in solution to obtain *Janus* monomers (**10**) that can be directly used to assemble mixed sequence *Janus* PNA oligomers. α-(*S*-ethylazido)-N^{Cbz} aminoethylglycyl [α-(*S*-eaz)-N^{Cbz} *aeg*] (**8**), α-(*S*-ethyltriazolyl-C4-methyl-N9-A^{NHCbz})-N^{Cbz} aminoethylglycyl [α-(*S*-etz-N9-A^{NHCbz})-N^{Cbz} *aeg*] (**9**) and α-(*S*-ethyltriazolyl-C4-methyl-N9-A^{NHCbz})] aminoethylglycyl-thyminyl (T) *Janus* PNA monomers [α-(*S*-etz-N9-A^{NHCbz})-*aeg-T Janus* PNA] (**10**). In order to examine the potential of C_α-triazolyl linker for complementation of DNA, control PNA (**9**) with nucleobases linked only by triazole was synthesized using the precursor monomers (**8**) by post solid phase click reaction with the propargyl nucleobases (**11-13**).[14-17]

4. α-(*S*-eaz) *aeg* T   **5.** α-(*S-eaz*) *aeg* C   **6.** α-(*S*-eaz) *aeg* G   **7.** α-(*S*-eaz) *aeg* A PNA

8. α-(*S*-eaz)]-N^Cbz-*aeg*   **9.** α-(*S*-etz -N9-A^NHCbz)-N^Cbz *aeg*   **10.** α-(*S*-etz-N9-A^NHCbz)-*aeg*-T

11   **12**   **13**
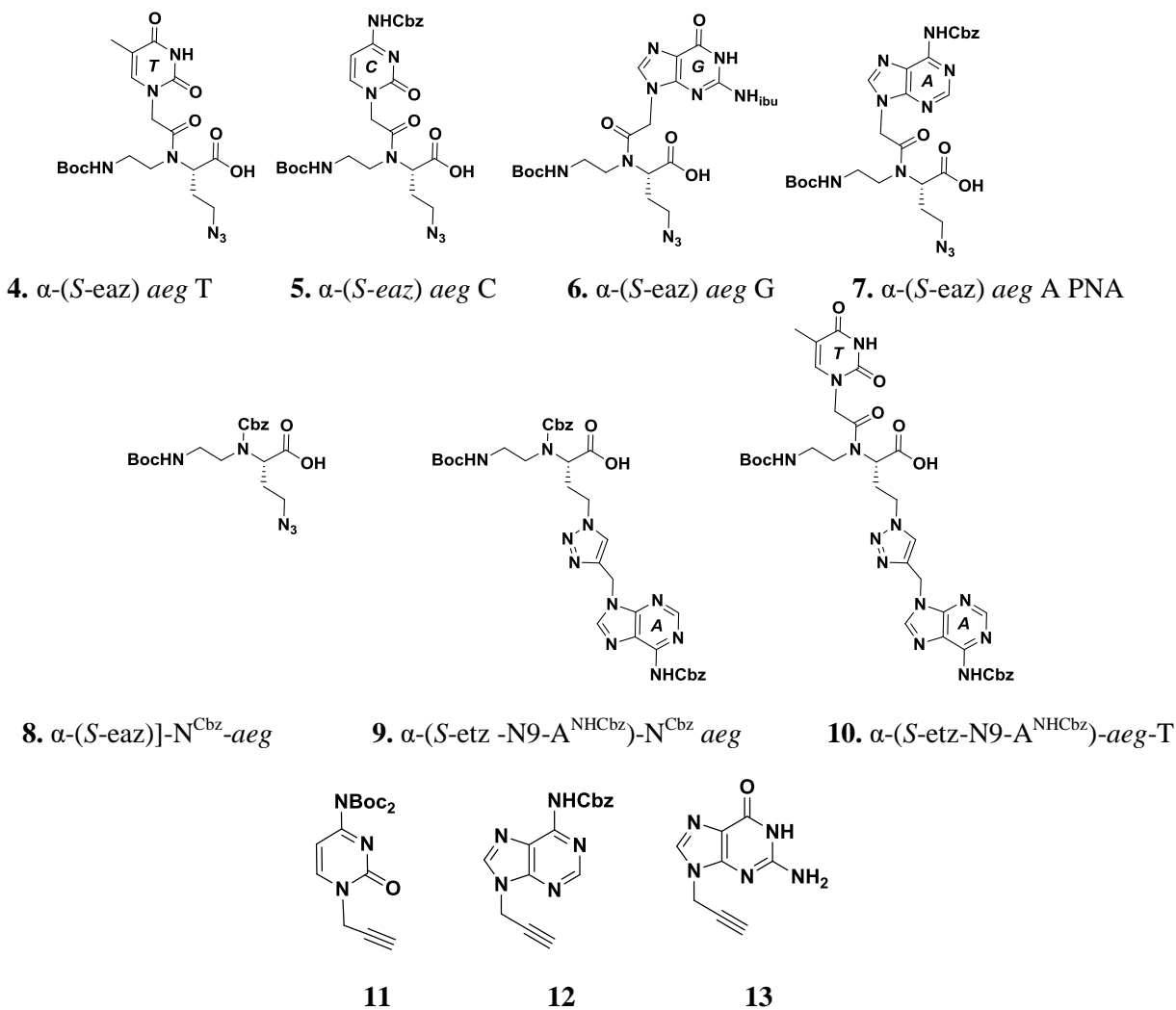
Figure 2.6 Various Modified PNA monomers and their precursors used in present study

2.2 Objectives of the present work

The specific objectives of this section are

➢ Synthesis of C$_α$ (*S*-ethylazido) *aeg* PNA A/T/C/G monomers (**4-7**)

➢ Synthesis of α-(*S*-ethylazido) N^Cbz *aeg* and C$_α$ (*S*-ethyltriazolyl-N9-A^NHCbz) *aeg* *N*^cbz monomers (**8-9**)

➢ Synthesis of α-(*S*-ethyltriazolyl-C4-methyl-N9- A^NHCbz)] *aeg* T *Janus* monomers (**10**)

➢ Synthesis of N1-propargyl C and N9-propargyl G/A (**11**-**13**)

➢ Characterization of rationally synthesized PNA monomers and their intermediates by various spectroscopic techniques

2.3 Synthesis of modified PNA monomers

This section describes the synthesis of rationally designed C_α (*S-eaz*) *aeg* PNA A/T/G/C monomers **4**-**10.**

2.3.1 Synthesis of C_α *(S-eaz) aeg A/T/C/G and* C_α *(S-eaz)-*N^{Cbz} *aeg monomers*

The commercially available L-glutamine **14** was treated with di-*t*-butyl dicarbonate [(Boc)2O] to obtain NH-Boc-L-Glutamine **15** (Scheme 2.1). The reaction of **15** with iodobenzene diacetate (PIDA) led to the conversion of its amido group to amino function to give 4-amino-2-(NHBoc) butanoic acid **16**.[18,19] This was followed by the diazo transfer reaction[20] in which the primary amine **16** was transformed to azide **17** by treatment with triflic anhydride and sodium azide (in-situ generation of triflyl azide) in presence of CuSO4.5H2O.[20] Compound **17** was converted to methyl ester derivative **18** using dimethyl sulfate followed by treatment with NHBoc-aminoacetaldehyde[21] to obtain C_α–substituted aminoethyl glycyl derivative **19.** The secondary amino group in **19** was protected as Cbz group[22] to get compound **20** that was hydrolyzed under basic conditions to afford the desired NHBoc-*eaz*-N^{CBz} monomer **8** in very good yield. This monomer was used for synthesis of control PNA which has nucleobases only on the triazole face (Scheme 2.1)

Scheme 2.1. Synthesis of C_α *(S-eaz)* N^{cbz} *aeg monomer (8)*
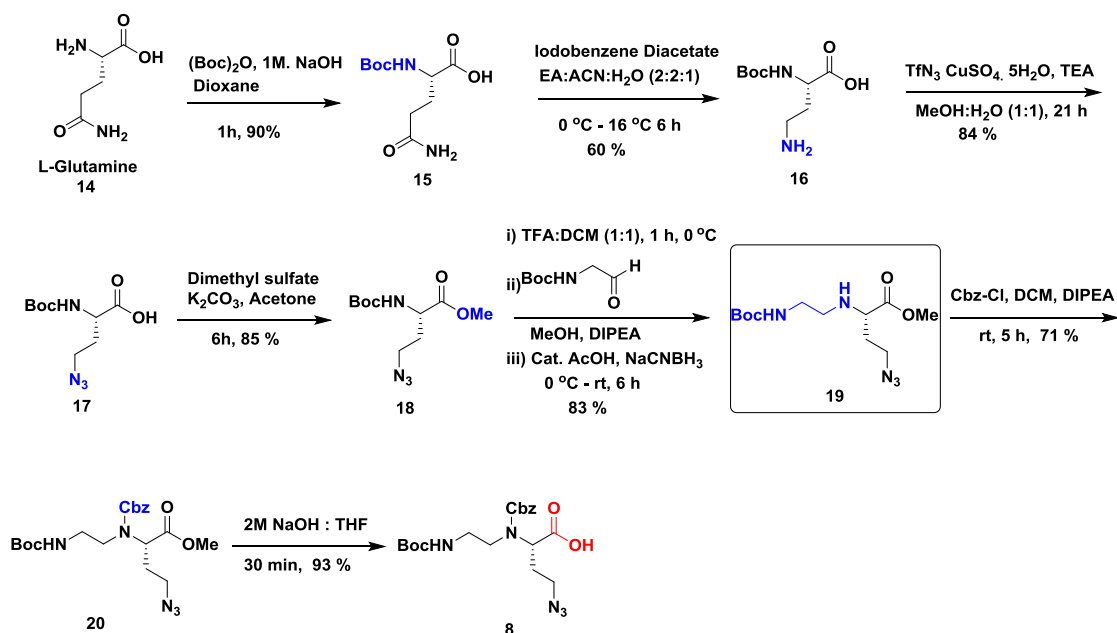
Compound **19** was coupled with pyrimidine (C/T) N1-acetic acids[23] using coupling agents DhBtOH and DCC (Scheme 2.2) to obtain the esters **21** and **22**, to get better yield of the corresponding ester monomers. These were hydrolyzed under basic conditions to afford the desired *aeg* PNA monomers bearing C_α-ethylazido substitution **4** and **5** respectively. When this procedure using coupling agents DhBtOH and DCC was tried with N-9 purine (A/G) acetic acids to make the corresponding monomers, the reaction unfortunately failed. Hence coupling of **19** with purine acetic acids derived from **A** and **G** was attempted using N-methylmorpholine (NMM) and isobutyl chloroformate (IBCF)[24] to give the esters **23** and **24**. These were hydrolyzed under basic conditions to afford the desired Boc protected PNA monomer acids with C_α-ethylazido substitution **6** and **7** respectively (Scheme 2.2).

Scheme 2.2. Synthesis of C_α (S-eaz) aeg C/T/G/A PNA monomers (4-7)
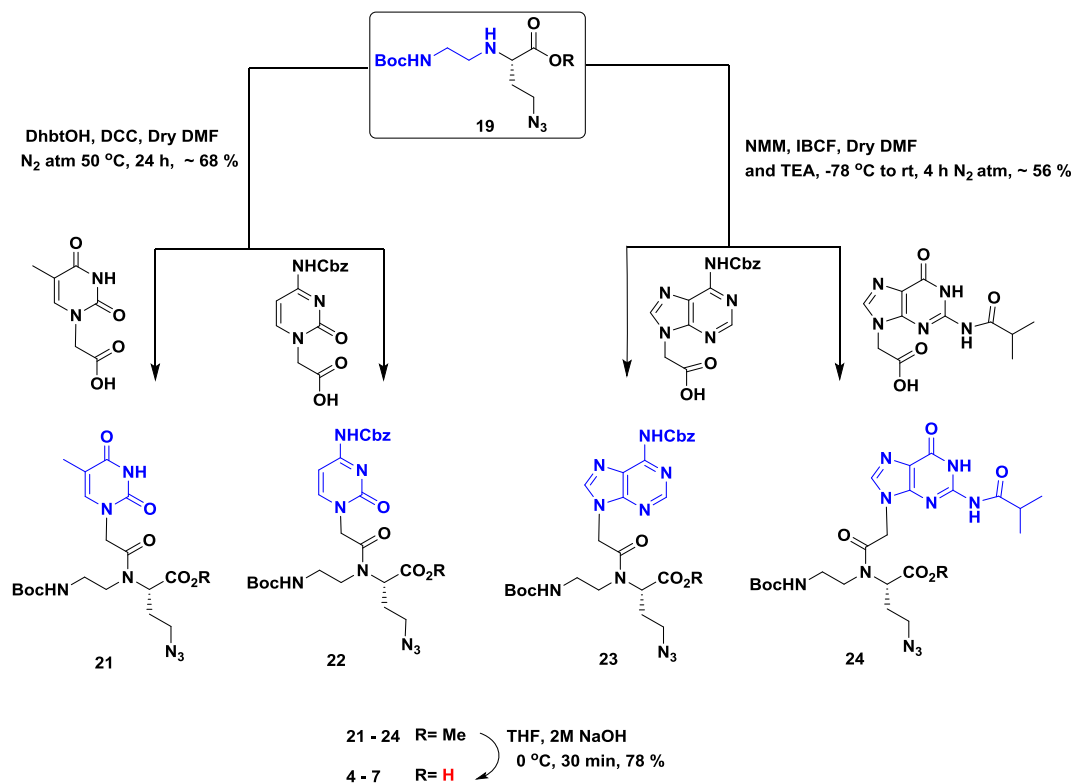
2.3.2 Synthesis of C_α(S-etz)-N9(A^{NHCbz})-NH^{Cbz}aeg (9) and [C_α(S-etz)-N9(A^{NHCbz})]-aeg-T (10) monomer

The azide **19** was reacted with N9 (propargyl) A^{NHCbz} **12** under click reaction conditions to get the block derivative ester **25** that was hydrolyzed under basic conditions to afford the desired trizolyl PNA monomer **9.** Similarly, compound **21** was reacted with N9(propargyl)-

A[NHCbz] **12** under click reaction condition to get *Janus* monomer ester **26**, which was hydrolyzed under basic conditions[19] to afford the desired *Janus* monomer **10** (Scheme 2.3)

Scheme 2.3. Synthesis *of NH[Boc] [C$_\alpha$ (S-etz-N9-A[NHCbz])] aeg T Janus monomer*
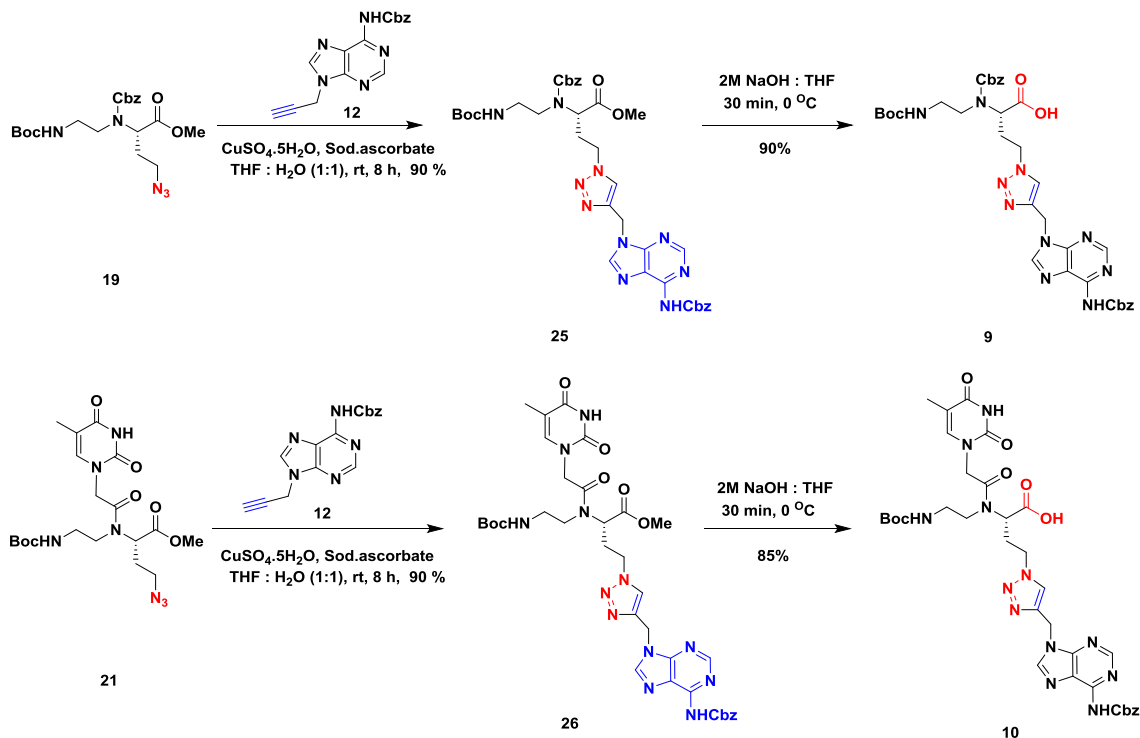
2.3.3 Synthesis of *NHBoc-N[Cbz]-aeg monomers*

The commercially available ethylenediamine **27** was treated with di-*t*-butyl dicarbonate [(Boc)2O] to get mono *NH*-Boc-protected ethylenediamine **28** (Scheme 2.4). This was N-alkylated by the reaction with ethylbromoacetate, triethylamine in ACN solvent to form compound **29.**[25] Compound **29** was then protected with Cbz group in presence of DIPEA in dry DCM as reported in to obtain Compound **30**, which were further hydrolyzed under basic conditions to afford the desired BocNH protected monomer **31** in very good yield for synthesis of control aeg oligomer.

Scheme 2.4 Synthesis of *NHBoc-N^{cbz} aeg monomer*
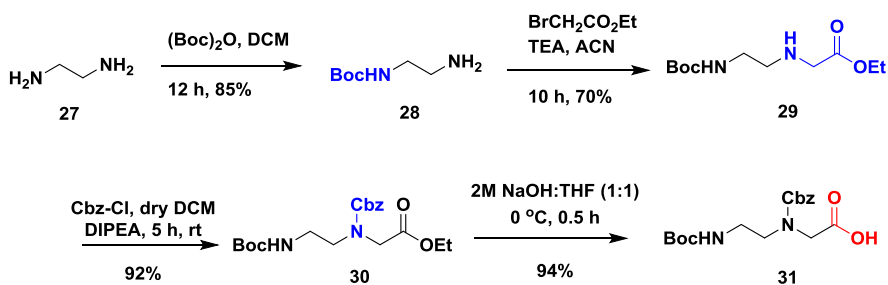
2.3.4 Synthesis of nucleobase acetic acid

This section describes the synthesis of nucleobase (A/G/C/T)-acetic acid from the corresponding nucleobases.

2.3.4a Synthesis of *(N^4-Benzyloxycarbonyl)-cytidinyl-acetic acid*. The commercial available cytosine **32** was alkylated with ethyl bromoacetate in the presence of NaH obtain compound **33** followed by protection of 4-NH$_2$ with Cbz group in presence of DMAP to obtained **34**, where ester group hydrolyzed in presence of basic medium to afford cytosine acetic acid **35** in 43% overall yield[26] (Scheme 2.5).

Scheme 2.5 Synthesis of *(N^4-Benzyloxycarbonyl)-cytidinyl-N1-acetic acid*
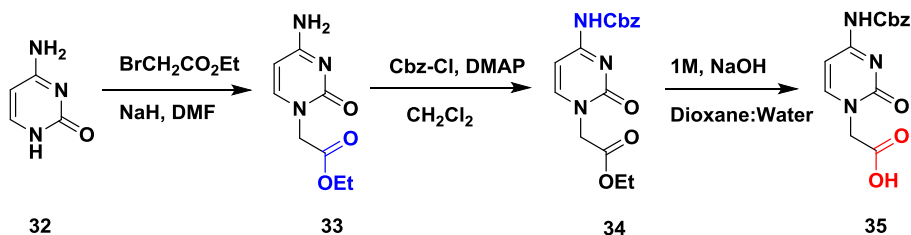
2.3.4b Synthesis of *N^6(Benzyloxycarbonyl)-adeninyl-N9-acetic acid*. The synthesis was started from compound **36** in which 6-NH$_2$ group of adenine was protected with Cbz-Cl to get compound **37** followed by alkylation by tertiary butyl to get compound **38** in the next step tertiary butyl group was undergo cleavage to afford the desired N9 derivative acid **39**.[27] (Scheme 2.6)

Scheme 2.6 Synthesis of N^6-(Benzyloxycarbonyl)-adeninyl-N9-acetic acid
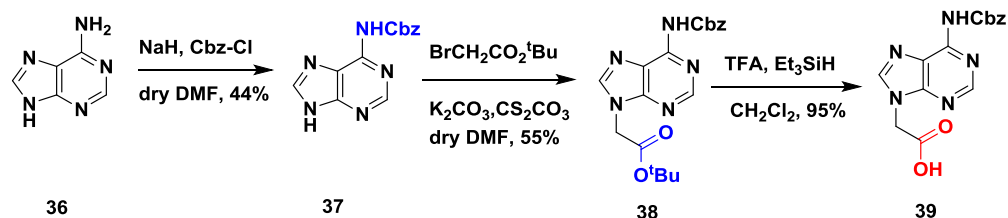
2.3.4c **Synthesis of *N^2-(Isobutanoyl)-guaninyl-N9-acetic acid*** - In the first step, the 2-NH$_2$ group of guanine **40** was protected in presence of isobutyric anhydride and dry DMF to get compound **41**, in next step selectively N9 of guanine alkylated in presence of tertiary butyl bromoacetate to get compound **42**, again tertiary butyl group cleavage in presence of triethylsilane and DCM to get desired N9 derivative compound **43.**[28-33] (Scheme 2.7)

Scheme 2.7 Synthesis of *N^2-(Isobutanoyl)-guaninyl-N9-acetic acid*
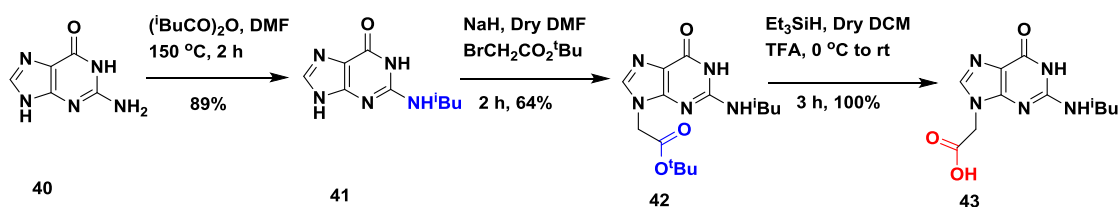
2.3.4d **Synthesis of *thymin-N1-yl acetic acid*** - alkylation of compound **44** in presence of ethylbromo acetate, by in situ hydrolysis in basic medium to get the desired compound **45.**[20,] (Scheme 2.8)

Scheme 2.8 Synthesis of *thymin-N1-yl acetic acid*
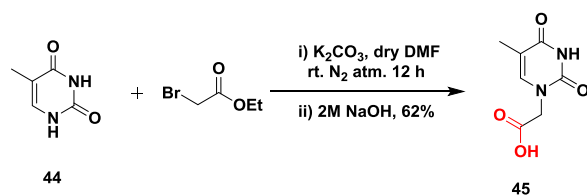
2.3.5 Synthesis of N-propargyl nucleobases

Cytidine (**46**) was protected in the presence of excess di-*t*-butyl dicarbonate [(Boc)$_2$O] to obtain the N^1,N^4,N^4-tris-Boc cytidine **47**, which was converted into N^4,N^4-di-boc **48** in the presence of NaHCO$_3$ in MeOH. This was then subjected to propargylation using NaH to afford

the compound **11** (Scheme 2.9).[34] Adenine **37** was propargylated in presence of K_2CO_3 in DMF to get compound **12** in good amounts (Scheme 2.10). 6-Chloro-2-aminopurine **49** was propargylated in the presence of K_2CO_3 to obtain the propargyl derivative of chloropurine **50** which was then converted to N9(propargyl)guanine **13.** (Scheme 2.11).

Scheme 2.9 Synthesis of *Di(tert-butyl)[2-(Prop-2-ynoyl)pyrimidin-4-yl]imidodicarbonate*
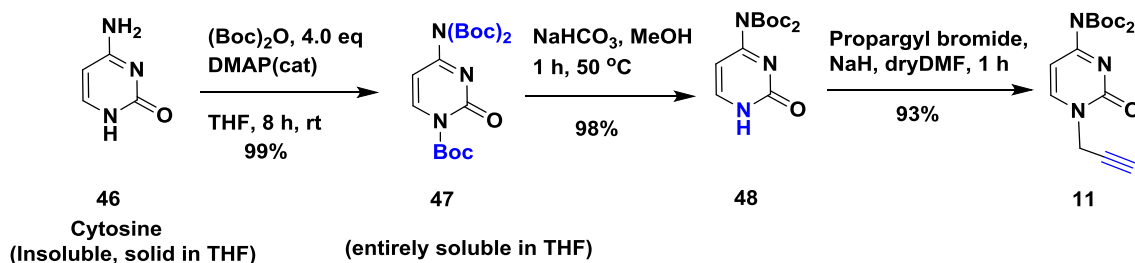
Scheme 2.10 Synthesis of *N-9-Propargyl adenine*
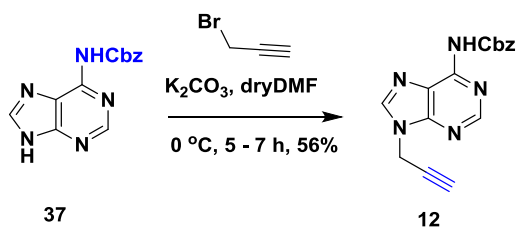
Scheme 2.11 Synthesis of *N-9-Propargylguanine*
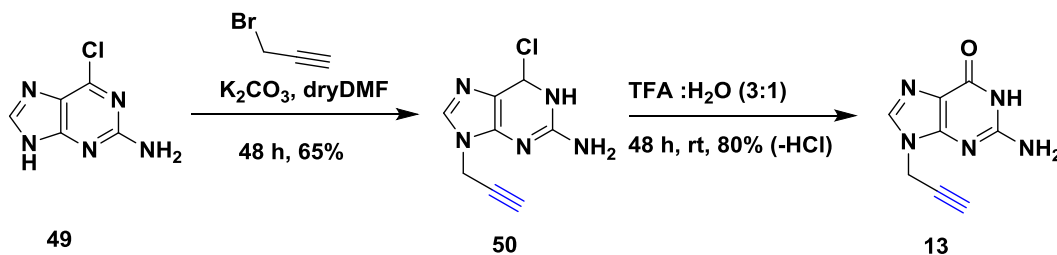
2.4 Summary

In summary, this section described the synthesis and characterization of rationally designed C_α-substituted PNA monomers, synthesis of propargyl derivative linked nucleobases and click reaction with the azide to get triazole. These monomers can have nucleobases on two faces, and so can be used for synthesis of the *Janus* PNA oligomers. All the intermediates have been characterized by ^{1}H & ^{13}C NMR spectroscopy and mass spectral analysis section 2.5. The

next chapter deals with the incorporation of these modified monomers into *Janu*s PNA oligomers using solid phase peptide synthesis.

2.5 Experimental section

This section details the synthetic procedures and spectral characterization of various rationally designed PNA monomers.

2.5.1 General

The chemicals used were of laboratory or analytical grade. All the solvents used were distilled or dried to carry out different reactions. Reactions were monitored by thin layer chromatography (TLC). Usual workup involved sequential washing of the organic extract with water and brine followed by drying the organic layer over anhydrous sodium sulphate and evaporation of solvent under vacuum. TLCs were carried out on pre-coated silica gel GF_{254} sheets (Merck 5554). TLCs were analysed under UV lamp, by iodine spray and by spraying with ninhydrin solution, followed by heating of the plate. Column chromatographic separations were performed using silica gel (60-120 or 100-200 mesh).

^{1}H and ^{13}C NMR spectra were recorded using Bruker AC-400 (400 MHz) or JEOL 400 MHz NMR spectrometers. The delta (δ) values for chemical shifts are reported in ppm and are referred to internal standard TMS or deuterated NMR solvents. The optical rotation values were obtained on Rudolph Research Analytical Autopol V polarimeter. Mass spectra for reaction intermediates were obtained by, Synapt G2 High Definition Mass Spectrometry.

2.5.2 Experimental procedures and spectral data

4-Carboxyamido-2(*S*)-amino(NH-t-butoxycarbonyl) pentanoic acid (15)
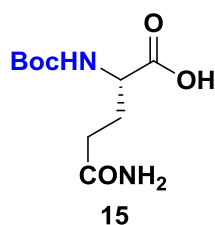
A solution of di-*tert*-butyl dicarbonate [(Boc)$_2$O] (16.4 g, 17.26 ml, 75.3 mmol) in dioxane (100 ml) was added in portion to an ice-cold solution of L-glutamine **14** (10 g, 68.5 mmol) in aq. NaOH (1 N, 100 ml). The reaction mixture was further stirred at 0 °C for 1 h. After completion of the reaction, dioxane was removed completely under vacuum from the reaction mixture. The aqueous layer was washed with diethyl ether to remove excess (Boc)$_2$O. The aqueous layer was cooled in an ice-water bath, acidified to pH 2-3 by slow addition of saturated KHSO$_4$

solution and then extracted with EtOAc (3×150 ml). The combined organic extracts were dried over anhydrous Na_2SO_4, filtered and concentrated to give compound **15** as a white powder, which was used without further purification (14.5 g, 86% yield). mp = 119-121 $^{\circ}$C; R_f = 0.39 in EtOAc/MeOH (50:50); $[\alpha]^{25}_D$ - 2.960 (*c* 0.5, MeOH); MS (MALDI-TOF) *m/z* calcd for $C_{10}H_{18}N_2O_5$ [M + K]$^+$ 285.0853, found 285.0501.

4-Amino-2(*S*)-amino(NH-t-butoxycarbonyl) butanoic acid (16) [13]
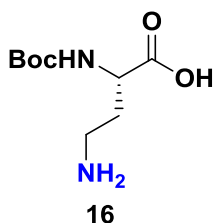
A slurry of compound **15** (5 g, 20.3 mmol) in EtOAc:ACN (1:1, 48 ml) and water (12 ml) and iodobenzene diacetate (7.87 g, 24.4 mmol) was cooled and stirred at 16 $^{\circ}$C for 30 min. The temperature was allowed to reach 20 ^{o}C and the reaction mixture was stirred until completion (approximately 4 h). Then, the reaction mixture was cooled to 0 ^{o}C and filtered under vacuum. The filter cake was washed with EtOAc and dried in vacuum to obtain compound **16** (2.65 g, 59% yield). mp = 200-201 $^{\circ}$C; R_f = 0.2 EtOAc/MeOH (50:50); $[\alpha]^{25}_D$ + 13.6 (*c* 0.5, MeOH); MS (MALDI-TOF) *m/z* calcd for $C_9H_{18}N_{42}O_4$ [M + K]$^+$ 257.0904, found 257.0740.

4-Azido-2(*S*)-amino(NH-t-butoxycarbonyl) butanoic acid (17) [14]
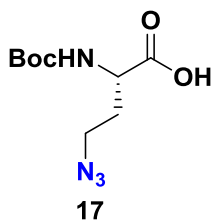
Sodium azide (13.3 g, 200 mmol) was dissolved in DCM (72 ml) and water (36 mL) and cooled to 0 ºC. Triflic anhydride (6.9 mL, 40.9 mmol) was added dropwise and left to stir for 2 h, after which the reaction was allowed to warm to room temperature and stirred for a further 3 h period. The layers were separated and the aqueous layer was extracted with DCM (2 x 60 mL). The combined organic layers were washed with aq. Na_2CO_3 (2 x 80 ml), further the organic layer was added a solution of compound **16** (4.46 g, 20.5 mmol), copper (II) sulphate hydrate (50 mg, 0.2 mmol) and potassium carbonate (4.2 g, 30.7 mmol) in water (75 ml) and MeOH (110 ml). After 21 h, TLC (water:acetic acid:n-butanol, 1:1:4) indicated the formation of a product (R_f = 0.3) and complete consumption of starting material (R_f 0.8). The organic were removed by evaporation, and the remaining aqueous fraction was extracted with DCM (2 x 150 ml). The aqueous layer was diluted with water (80 ml) and acidified to pH 2.0 with aq. HCl (1 M) and extracted with DCM (3 x 150 ml). The combined organic layers were washed with aq.

sat. brine (200 ml), dried (MgSO$_4$), filtered and concentrated *in vacuo* to afford the desired product of 2(*S*)-amino(NH-t-butoxycarbonyl)-4-azidobutanoic acid **17** (4.16 g, 84%) as a colourless oil.

4-Azido-2(*S*)-amino(NH-*t*-butoxycarbonyl) methyl butanoate (18)
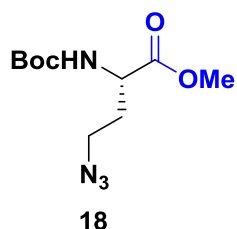
To a stirred solution of compound **17** (5.90 g, 24.2 mmol) and K$_2$CO$_3$ (8.35 g, 60 mmol) in dry acetone (100 ml) was added dimethyl sulfate (2.7 ml, 21.8 mmol) and the reaction mixture was heated to 55 °C for 4 h under reflux condenser. Acetone was evaporated completely and water (100 ml) was added to the concentrate, which was then extracted with EtOAc (3 x 80 ml). The combined organic layer was washed with brine, dried over an. Na$_2$SO$_4$, filtered and concentrated. The residue was purified on silica gel (60-120 mesh) using petroleum ether (pet ether) and EtOAc to give compound **18** as colorless liquid (5.3 g, 85% yield); R$_f$ = 0.5 pet ether/EtOAc (70:30); [α]^{25} $_D$ - 17.8 (*c* 1, MeOH); IR (neat) 3358, 2975, 2098, 1702, 1511, 1445, 1359, 1247, 1159, 1054, 859, 775, 636, 557 cm^{-1}; ^{1}H NMR (400 MHz, CDCl$_3$) δ 5.22 (s, 1H), 4.41 (d, *J* = 4.1 Hz, 1H), 3.77 (d, *J* = 0.8 Hz, 3H), 3.41 (t, *J* = 6.6 Hz, 2H), 2.12 (dd, *J* = 13.1, 6.1 Hz, 1H), 1.93 (dq, *J* = 13.7, 6.7 Hz, 1H), 1.45 ppm (s, 9H); ^{13}C NMR (100 MHz, CDCl$_3$) δ 172.5, 155.3, 80.1, 58.6, 52.5, 51.3, 47.7, 31.8, 28.3 ppm; HRMS (ESI-TOF) m/z calcd for C$_{10}$H$_{18}$N$_4$O$_4$ [M + Na]$^+$ 281.1225, found 281.1218.

4-Azido-[2(*S*)-amino(N-ethylamino-t-butoxycarbonyl)] methyl butanoate (19)
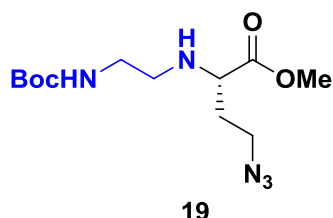
Compound **18** (2.65 g, 10.26 mmol) was dissolved in DCM (5.3 mL) and TFA (5.3 ml) was added to it at 0 °C. The mixture was stirred for 1 h. Toluene (25 ml) was added and the volatiles were removed under vacuum to obtain TFA salt of deprotected amine of **18** (4.1 g, 15.18 mmol) as crude oil. Without further purification, this was dissolved in MeOH (50 mL) to which N-Boc-aminoacetaldehyde (1.80 g, 11.4 mmol) in MeOH (15 ml) was added, while continuous stirring at 0 °C. DIPEA (5.40 ml, 42.0 mmol) was slowly added dropwise to the stirred solution. After stirring for 30 min, glacial acetic acid (1.2 ml) and NaCNBH$_3$ (1.4 g, 22.8 mmol) were sequentially added and the reaction mixture was stirred for 3 h. All volatiles were removed under vacuum and the residue was dissolved in EtOAc and extracted with saturated aq. NaHCO$_3$. The organic phase, after drying with MgSO$_4$, was

removed under vacuum and the residue was purified by flash chromatography eluting with EtOAc -pet ether. Compound **19** was obtained as yellow oil (2.6 gm, 83.2 % yield) $[\alpha]^{25}_{D}$ - 15.6 (*c* 1, MeOH); IR (neat) 3348, 2974, 2098, 1698, 1513, 1450, 1363, 1250, 1161, 1051, 990, 863, 777, 740, 638, 557 cm^{-1}; ^{1}H NMR (400 MHz, CDCl$_3$) δ 5.04 (s, 1H), 3.75 (d, *J* = 4.7 Hz, 3H), 3.54 – 3.38 (m, 2H), 3.37 – 3.32 (m, 1H), 3.23 (s, 1H), 3.19 3.11 (m, 1H), 2.80 (ddt, *J* = 9.5, 6.7, 4.8 Hz, 1H), 2.60 – 2.49 (m, 1H), 1.94 (ddd, *J* = 5.1, 4.0, 1.9 Hz, 1H), 1.78 (tdd, 10.0, 5.2, 3.0 Hz, 2H), 1.46 – 1.45 ppm (m, 9H); ^{13}C NMR (100 MHz, CDCl$_3$) δ 175.2, 156.1, 79.1, 58.3, 52.0, 48.1, 47.5, 40.3, 32.5, 28.4 ppm; HRMS (ESI-TOF) m/z calcd for C$_{12}$H$_{23}$N$_5$O$_4$ [M + H]$^+$ 302.1828, found 302.1832.

4-Azido-[2(S)-amino(N-benzyloxycarbonyl-N-ethylamino-t-butoxycarbonyl)] methyl butanoate (20)
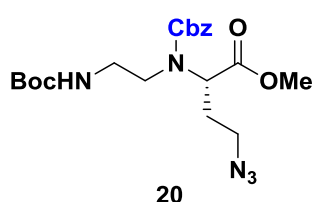
To a solution of **19** (0.65 g, 2.17 mmol) in DCM (15 ml) at 0 °C was added benzyl chloroformate (0.75 g, 4.35 mmol) dropwise followed by addition of N,N-diisopropylethylamine (1.2 g, 8.7 mmol). After stirring at room temperature for 6 h, saturated aq. NH$_4$Cl solution (30 ml) was added and extracted with DCM (4 x 30 ml). The organic layers were dried over MgSO$_4$, concentrated *in vacuo* and purified with flash chromatography (Eluent: 50-60% EtOAc in hexane) to afford compound **20** (0.67 g, 71%) as yellow oil; $[\alpha]^{25}_{D}$ - 66.4 (*c* 1, MeOH); IR (neat) 3887, 3735, 3388, 2973, 2100, 1697, 1509, 1455, 1362, 1246, 1162, 1084, 1046, 1002, 864, 735, 699, 607 cm^{-1}; ^{1}H NMR (400 MHz, CDCl$_3$, 25 °C, TMS) δ = 7.35 – 7.3 (m, 5H), 5.21 – 5.04 (m, 3H), 4.15 – 4.09 (m, 1H), 3.76 – 3.69 (m, 2H), 3.54 – 3.45 (m, 2H), 3.35 – 3.28 (m, 3H), 3.21 – 3.15 (m, 1H), 2.36 – 2.17 (m, 2H), 2.11 – 1.8 (m, 1H), 1.42 ppm (s, 9H); ^{13}C NMR (100 MHz, CDCl$_3$, 25 °C, TMS) δ = 171.6, 171.4, 156.0, 155.9, 136.2, 136.0, 128.7, 128.4, 128.3, 127.9, 79.4, 67.8, 58.9, 58.3, 52.8, 52.6, 48.9, 48.7, 48.5, 48.3, 39.6, 29.9, 28.9, 28.5 ppm; HRMS (ESI-TOF) m/z calcd for C$_{20}$H$_{29}$N$_5$NaO6 [M + Na]$^+$ 458.2016, found 458.2018.

4-Azido-[2(*S*)-amino(N-benzyloxycarbonyl-N-ethylamino-t-butoxycarbonyl)] butanoic acid (8)
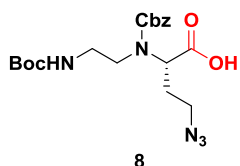
To a stirred solution of compound **20** (0.65 g, 1.55 mmol) in THF (12 mL) was added aq. NaOH (2 M, 12 ml) at 0 °C, and the reaction mixture was stirred for 30 min at 0 °C. Upon completion of the reaction, as confirmed by

TLC, water (10 mL) was added, and the resulting mixture was extracted with EtOAc (2 × 30 ml). The aqueous layer was neutralized with saturated aq. KHSO$_4$ at 0 °C to pH 3 - 4 and then extracted with EtOAc (4 × 30 ml). The combined organic extracts were dried over anhydrous Na$_2$SO$_4$. The solvent was evaporated in vacuo, and crude product was purified by column chromatography (Eluent: 60–70% EtOAc in hexane) to afford monomer **8** (0.6 g, 91%) as a hygroscopic solid. $[\alpha]^{25}_D$ - 42.0 (*c* 1, MeOH); IR (neat) 3825, 3736, 3359, 2977, 2313, 2100, 1696, 1513, 1459, 1420, 1364, 1247, 1159, 1083, 1047.10, 858, 771, 734, 698, 561, 531 cm^{-1}; ^{1}H NMR (400 MHz, DMSO-d$_6$, 25°C) δ = 7.37 – 7.29 (m, 5H), 6.87 – 6.51 (m, 1H), 5.11 – 4.97 (m, 2H), 4.27 - 4.23 (dd, *J* = 9.0, 5.0 Hz, 1H), 3.49 – 3.33 (m, 3H), 3.12 – 2.95 (m, 3H), 2.19 – 1.92 (m, 2H), 1.37 ppm (s, 9H); ^{13}C NMR (100 MHz, DMSO-d$_6$, 25°C) δ = 172.3, 155.5, 136.7, 128.3, 127.7, 126.6, 126.4, 79.2, 66.3, 62.8, 58.1, 48.1, 46.9, 28.4, 28.1 ppm, HRMS (ESI-TOF) m/z calcd for C$_{19}$H$_{28}$N$_5$O$_6$ [M + H]$^+$ 422.2040, found 422.2045.

### 4-Azido-[2(*S*)-amino(N-ethylamino-t-butoxycarbonyl)-N-acetamido(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)] methyl butanoate (21)

21
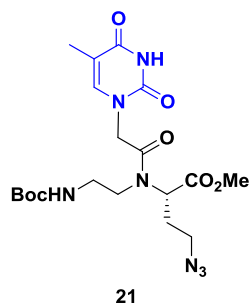
To a stirred solution of thymine-1-ylacetic acid **45,** (1.0 g, 5.3 mmol) in dried DMF (50 ml) under an inert atmosphere were added DCC (1.1 g, 5.3 mmol) and DhbtOH (0.9 g, 5.3 mmol). The resulting mixture was stirred at room temperature for 1 h. Compound **19** (1.5 g, 4.8 mmol) in dry DMF (30 ml) was added to the above reaction mixture, and stirred at 50 °C for 24 h. The solvent was removed under reduced pressure, and the remaining residue was dissolved in EtOAc (300 ml) and washed with a saturated aq. NaHCO$_3$ (300 ml) followed by 10% aq. KHSO$_4$ (300 ml). The organic layer was washed with brine (3 × 150 ml) and dried over Na$_2$SO$_4$. The solvent was removed under reduced pressure. The crude product was purified by column chromatography to afford a **21** (1.45 g, 70%) as a white solid; m.p. 80-83 °C; $[\alpha]^{25}_D$ - 14.0 (*c* 1, MeOH); IR (neat) 3732, 3379, 3185, 3057, 2977, 2102, 1667, 1509, 1457, 1366, 1247, 1163, 1095, 1006, 863, 782, 731, 640, 554 cm^{-1}; ^{1}H NMR (400 MHz, CDCl$_3$) δ 9.13 (s, 1H), 7.04 – 6.92 (m, 1H), 5.61 (s, 1H), 4.80 (t, *J* = 16.2 Hz, 1H), 4.18 (d, *J* = 16.2 Hz, 1H), 3.93 – 3.89 (m, 1H), 3.81-3.07 (m, 4H), 3.53 (dt, *J* = 12.5, 5.0 Hz, 2H), 3.34 (ddd, *J* = 14.0, 7.5, 3.8 Hz, 2H), 3.29 – 3.17 (m, 1H), 2.31 (tt, *J* = 14.1, 7.0 Hz, 2H), 1.94 – 1.91 (m, 3H), 1.45 ppm (s, 9H); ^{13}C NMR (100 MHz, CDCl$_3$) δ 171.2, 167.5, 164.2, 156.1, 151.1, 141.0,

110.9, 80.1, 59.0, 53.0, 49.2, 48.6, 48.3, 39.0, 34.0, 28.5, 28.3, 12.5 ppm; HRMS (ESI-TOF) m/z calcd for $C_{19}H_{29}N_7O_7$ [M + Na]$^+$ 490.2025, found 490.2026.

4-Azido-[2(S)-amino(N-ethylamino-t-butoxycarbonyl)-N-acetamido(2-oxopyrimidin-1(2H)-yl)-N^4(benzyloxycarbonyl)] methyl butanoate (22)
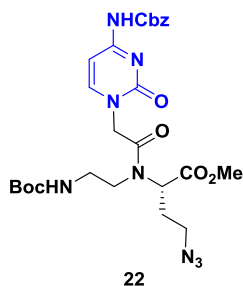
By following the same procedure described for **21,** compound **22** was prepared, from **19** and N4-benzyoxycarbonyl cytidinyl-N1-acetic acid **35**, (67 %), purified (Eluent: 80 - 90 % EtOAc in petroleum ether), m.p. 75-78 °C; $[\alpha]^{25}_D$ – 23.2 (*c* 1, MeOH); IR (neat) 3254, 2972, 2101,1740, 1659, 1557, 1498, 1447, 1366, 1193, 1063, 997, 853, 790, 736, 698, 629, 573 cm^{-1}; ^{1}H NMR (400 MHz, CDCl$_3$, 25 °C, TMS) δ = 7.83 (s, 1H), 7.56 (dd, *J* = 23.9, 6.2 Hz, 1H), 7.38 (d, *J* = 3.0 Hz, 5H), 7.24 (d, *J* = 7.3 Hz, 1H), 5.57 (s, 1H), 5.25 – 5.17 (m, 2H), 5.03 (d, *J* = 15.4 Hz, 1H), 4.25 (dd, *J* = 14.2, 7.0 Hz, 1H), 3.96 – 3.78 (m, 2H), 3.75 (s, 3H), 3.52 (dt, *J* = 12.5, 4.9 Hz, 2H), 3.38 (dt, *J* = 19.0, 5.8 Hz, 2H), 3.25 (d, *J* = 5.4 Hz, 1H), 2.29 (dd, *J* = 12.6, 7.0 Hz, 2H), 1.43 (d, *J* = 8.6 Hz, 9H); ^{13}C NMR (100 MHz, CDCl$_3$, 25 °C, TMS) δ = 171.2, 167.2, 162.9, 156.0, 155.6,152.2, 149.6, 134.93, 128.7, 128.3, 95.0, 80.0, 68.0, 59.0, 52.8, 50.3, 49.5, 48.1, 38.93, 28.5, 28.1, 20.3 ppm; HRMS (ESI-TOF) m/z calcd for $C_{26}H_{35}N_8O_8$ [M + H]$^+$ 587.2578, found 587.2581.

4-Azido-[2(*S*)-amino(N-ethylamino-t-butoxycarbonyl)-N-acetamido(5-methyl-2,4-dioxo-3,4-dihydropyrimi din-1(2H)-yl)] butanoic acid (4)
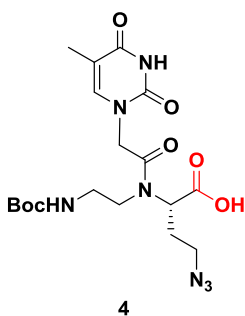
To a stirred solution of compound **21** (0.70 g, 1.55 mmol) in THF (15 mL) was added aq NaOH (2 M, 15 ml) at 0 °C, and the reaction mixture was stirred for 30 min at 0 °C. Upon completion of the reaction as confirmed by TLC, H$_2$O (50 ml) was added, and resulting mixture was extracted with EtOAc (2 × 35 ml). The aqueous layer was neutralized with saturated aq. KHSO$_4$ at 0 °C to pH 3 - 4 and then extracted with EtOAc (4 × 35 ml). The combined organic extracts were dried over anhydrous Na$_2$SO$_4$. The solvent was evaporated in vacuo, and crude product was purified by column chromatography to afford **4** as a white solid (0.63 g, 89%), m.p. 132-135°C; $[\alpha]^{25}_D$ -22.6 (*c* 1, MeOH); IR (neat) 3825, 3733, 3186, 3058, 2979, 2312, 2101, 1662, 1511, 1461, 1365, 1246, 1162, 1096, 1013, 964, 857, 780, 732, 637, 555 cm^{-1}; ^{1}H NMR (400 MHz, DMSO-d$_6$) δ 11.30 (s, 1H), 7.36 (s, 1H), 6.91 (s, 1H),

4.70 − 4.45 (m, 2H), 4.42 − 4.14 (m, 1H), 3.56 − 3.2 (m, 6H), 3.13 − 2.87 (m, 1H), 2.34 − 1.86 (m, 2H), 1.75 (d, J = 0.7 Hz, 3H), 1.38 ppm (s, 9H); ^{13}C NMR (100 MHz, DMSO-d_6) δ 171.6, 168.0, 167.3, 164.4, 155.7, 151.0, 142.1, 108.0, 78.1, 57.6, 47.98, 47.92, 47.80, 47.1, 28.1, 11.9 ppm; HRMS (ESI-TOF) m/z calcd for $C_{18}H_{27}N_7O_7$ [M + Na]$^+$ 476.1869, found 476.1840.

4-Azido-[2(*S*)-amino(N-ethylamino-t-butoxycarbonyl)-N-acetamido(2-oxopyrimidin-1(2H)-yl)-N[4](benzyloxycarbonyl)] butanoic acid (5)
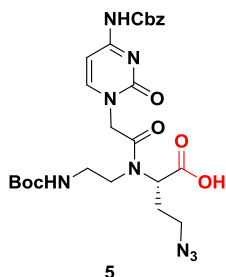
Cytosine acid monomer **5** (0.66 g, 1.15 mmol) was prepared, purified, and characterized the same way as described for **4**, starting from compound **22** (0.80 g, 1.36 mmol), to afford yield (85%), mp = 172 − 175 °C; $[α]^{25}_D$ − 24.80 (*c* 1, MeOH); IR (neat) 3387, 2974, 2100, 1656, 1561, 1502, 1452, 1369, 1206, 1066, 999, 789, 743, 696, 563 cm^{-1}; ^{1}H NMR (400 MHz, DMSO) δ 12.86 (s, 1H), 10.80 (s, 1H), 7.95 (d, J = 7.3 Hz, 1H), 7.45 − 7.33 (m, 5H), 7.02 (d, J = 7.3 Hz, 1H), 6.92 (s, 1H), 5.20 (s, 2H), 4.85 − 4.72 (m, 2H), 4.24 (dd, J = 8.8, 5.6 Hz, 1H), 3.58 − 3.25 (m, 6H), 2.18 − 1.97 (m, 2H), 1.38 (d, J = 8.7 Hz, 9H); ^{13}C NMR (101 MHz, DMSO) δ 171.5, 167.2, 163.2, 155.7, 155.0, 153.2, 150.9, 136.0, 128.5, 128.1, 127.9, 93.8, 78.1, 66.5, 57.6, 49.9, 47.8, 47.2, 28.2.ppm; HRMS (ESI-TOF) m/z calcd for $C_{25}H_{33}N_8O_8$ [M + H]$^+$ 573.2421, found 573.2424.

4-Azido-[2(*S*)-amino[N-ethylamino-t-butoxycarbonyl-N-acteamido(N[6]-benzyloxycarbonyl)-9H-purin-9-yl] methyl butanoate (23)
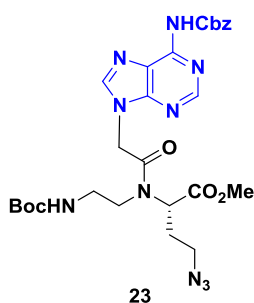
A solution of (0.8 g, 2.4 mmol) of adenine-9-yl-acetic acid **39** and (0.5 g 5.4 mmole) of N-methylmorpholine in freshly distilled DMF (15 ml) was cooled to 0 °C and was added 0.37 g (2.69 mmol) of isobutyl chloroformate in a single portion and stirred 5 - 7 min in an inert gas atmosphere. The temperature of reaction mixture was maintained at a -78 °C, to this 0.32 g (2.4 mmol) of DIPEA was added followed by a solution of 0.52 g (1.7 mmol) compound **19** in of distilled DMF (5 ml) and maintain the temperature -78 °C for next 20 min then for 3 h at room temperature, after which the solvent was distilled off and the residue was dissolved in EtOAc 30 mL. The organic layer was washed sequentially with saturated aq. NaCl solution (3 × 30 ml), saturated aq. NaHCO$_3$ solution (1 × 30 ml), and again with saturated aq. NaCl solution (3 × 30 ml). The organic layer was separated and dried over

Na$_2$SO$_4$. The solvent was distilled off and the residue was purified on silica gel (60 - 120 mesh) using DCM and MeOH to give compound **23,** 0.6 g (56%); mp = 60 – 63 °C; [α]^{25}$_D$ – 14.60 (*c* 1, MeOH); IR (neat) 3336, 2933, 2246, 2102, 1745, 1699, 1610, 1512, 1462, 1362, 1206, 1159, 1088, 987, 911, 859, 730, 641, 587 cm^{-1}; ^{1}H NMR (400 MHz, CDCl$_3$) δ 9.29 (d, *J* = 79.3 Hz, 1H), 8.77 – 8.66 (m, 1H), 8.05 – 7.95 (m, 1H), 7.34 (ddd, *J* = 18.7, 14.3, 6.4 Hz, 5H), 5.68 (s, 1H), 5.26 (t, *J* = 7.9 Hz, 2H), 5.14 (t, *J* = 14.3 Hz, 1H), 5.08 – 4.87 (m, 2H), 3.89 (ddd, *J* = 23.6, 8.7, 5.4 Hz, 2H), 3.80 – 3.64 (m, 3H), 3.63 – 3.15 (m, 4H), 2.31 (dq, *J* = 10.0, 5.2 Hz, 1H), 2.24 – 2.02 (m, 1H), 1.41 (d, *J* = 12.3 Hz, 9H); ^{13}C NMR (101 MHz, CDCl$_3$) includes all the major and minor rotamer, δ 170.8, 166.83, 166.5, 156.2, 153.1, 152.9, 151.6, 151.2, 149.6, 143.9, 135.6, 128.7, 128.6, 128.53, 127.0, 80.3, 67.7, 58.9, 53.2, 52.9, 49.4, 48.4, 44.1, 39.1, 28.5, 28.4, 28.3; HRMS (ESI-TOF) m/z calcd for C$_{27}$H$_{35}$N$_{10}$O$_7$ [M + H]$^+$ 611.2690, found 611.2702.

## 4-Azido-[2(*S*)-amino(N-ethylamino-t-butoxycarbonyl)-N-acetamido(N^2-isobutyramido-6-oxo-1,6-dihydro-9H-purin-9-yl)] methyl butanoate (24)
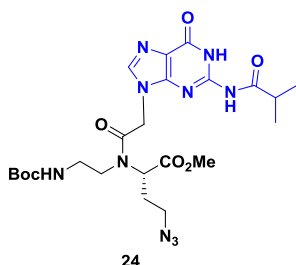
By following the same procedure described for **23,** compound **24** (0.73 g, 1.3 mmol) was prepared, from **19** (0.72 g, 2.39 mmol), and N^2-(isobutyryl)-N9-(carboxymethyl)-guanine **43,** to afford yield (54%), mp = 203 – 206 °C; IR (neat) 3850, 3734, 3392, 3345, 3186, 2966, 2101, 1681, 1611, 1558, 1469, 1411, 1372, 1251, 1203, 1028, 785, 580, 537 cm^{-1}; ^{1}H NMR (400 MHz, DMSO) δ 12.08 (s, 1H), 11.53 (s, 1H), 7.87 (s, 1H), 7.04 (s, 1H), 5.21 – 4.99 (m, 2H), 4.30 (dd, *J* = 14.3, 8.3 Hz, 1H), 3.61 (d, *J* = 13.1 Hz, 4H), 3.51 – 3.36 (m, 3H), 3.31 – 3.24 (m, 2H), 2.79 (dt, *J* = 13.7, 6.8 Hz, 1H), 2.17 (dd, *J* = 13.8, 6.3 Hz, 1H), 2.10 – 1.96 (m, 1H), 1.40 (s, 9H), 1.14 – 1.10 (d, 6H); ^{13}C NMR (101 MHz, DMSO -D6) δ 180.1, 170.4, 166.7, 155.8, 154.9, 147.9, 78.2, 57.7, 54.9, 52.2, 47.7, 47.2, 44.0, 34.7, 29.0, 28.2, 28.1, 18.9, 18.8 ppm; HRMS (ESI-TOF) m/z calcd for C$_{23}$H$_{35}$N$_{10}$O$_7$ [M + H]$^+$ 563.2690, found 563.2678.

4-Azido-[2(*S*)-amino(N-ethylamino-t-butoxycarbonyl)-N-acteamido(N^{6}benzyloxycarbonyl)-9H-purin-9-yl] butanoic acid (6)

Adenine acid monomer **6** (0.86 g, 1.44 mmol) was prepared, purified, and characterized the same way as described for **4**, starting from compound **23** (1.1 g, 1.8 mmol), to afford yield
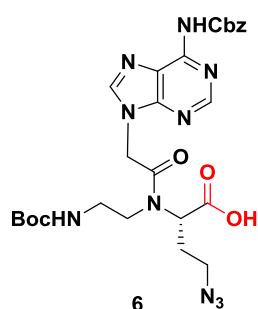
(80%), mp = 175 – 178 °C; $[\alpha]^{25}_D$ – 17.40 (*c* 1, MeOH); IR (neat) 3390, 2928, 2101, 1704, 1668, 1614, 1466, 1411, 1363, 1211, 1162, 1085, 991, 733, 638, 554 cm^{-1}; ^{1}H NMR (400 MHz, DMSO - D6) δ 10.67 (s, 1H), 8.60 (t, *J* = 8.7 Hz, 1H), 8.35 (d, *J* = 7.7 Hz, 1H), 7.39 (ddd, *J* = 23.3, 18.6, 6.8 Hz, 5H), 7.00 (d, *J* = 38.8 Hz, 1H), 5.37 (dd, *J* = 43.6, 27.3 Hz, 1H), 5.22 (s, 2H), 4.27 (d, *J* = 37.2 Hz, 1H), 4.06 – 3.79 (m, 1H), 3.76 – 3.55 (m, 2H), 3.40 (dd, *J* = 15.4, 8.0 Hz, 2H), 3.08 (dd, *J* = 34.2, 25.2 Hz, 1H), 2.26 (dd, *J* = 65.2, 5.8 Hz, 1H), 1.99 (s, 1H), 1.85 – 1.72 (m, 1H), 1.37 (d, *J* = 17.3 Hz, 9H); ^{13}C NMR (101 MHz, DMSO - D6) δ 167.8, 167.0, 156.2, 152.8, 151.9, 149.8, 145.7, 136.8, 128.9, 128.4, 128.30, 123.4, 79.6, 78.6, 78.0, 66.7, 58.4, 48.7, 44.6, 29.5, 28.6, 22.6 ppm; HRMS (ESI-TOF) m/z calcd for $C_{26}H_{33}N_{10}O_7$ [M + H]$^+$ 597.2534, found 597.2524.

4-Azido-[2(S)-amino(N-ethylamino-t-butoxycarbonyl-N-acetamido(N^2-isobutyramido-6-oxo-1,6-dihydro-9H-purin-9-yl] butanoic acid (7)
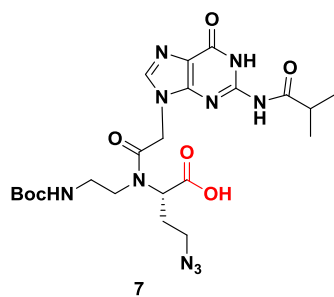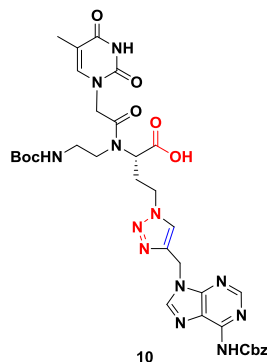
Guanine ester monomer **7** (0.65 g, 1.19 mmol) was prepared, purified, and characterized the same way as described for **4**, starting from compound **24** (0.9 g, 1.60 mmol), to afford yield (75%), IR (neat) 3739, 2929, 2311, 2261, 2103, 1683, 1645, 1591, 1528, 1175,866, 583 cm^{-1}; ^{1}H NMR (400 MHz, DMSO) δ 12.08 (s, 1H), 11.67 (s, minor 0.2H), 11.57 (s, major 0.8H), 7.96 (s, minor 0.2H), 7.87 (s, major 0.8H), 6.99 (s, major 0.7H), 6.87 (s, minor 0.3H), 5.20 – 4.85 (m, 3H), 4.64 (d, *J* = 5.5 Hz, minor 0.2H), 4.26 – 4.20 (m, major 0.8H), 3.56 (dd, *J* = 14.7, 7.0 Hz, 1H), 3.50 – 3.24 (m, 5H), 2.78 (dt, *J* = 12.0, 6.0 Hz, 1H), 2.16 (dd, *J* = 13.9, 6.2 Hz, 1H), 2.02 (dt, *J* = 14.0, 6.9 Hz, 1H), 1.38 (d, *J* = 11.1 Hz, 9H), 1.11 (dd, *J* = 6.7, 2.7 Hz, 6H); ^{13}C NMR (101 MHz, DMSO) δ 180.1, 171.5, 169.0, 166.6, 155.3, 149.2, 147.9, 140.6, 140.3, 119.5, 78.2, 57.7, 56.8, 48.0, 47.2, 44.4, 44.0, 34.7, 28.2, 18.9 ppm; HRMS (ESI-TOF) m/z calcd for $C_{22}H_{33}N_{10}O_7$ [M + H]$^+$ 549.2534, found 549.2540.

4-Triazolyl-[4-methyl-(N^6-benzyloxycarbonyl)-9H-purin-9-yl]-2(S)-[amino(N-ethylamino-t-butoxycarbonyl-N-benzyloxycarbonyl)] methyl butanoate (25)

To a mixture of **19** (0.25 g, 0.57 mmol) in THF (5.0 ml) and distilled water (5 ml) was added CuSO$_4$.5H$_2$O (10 mg, 0.03 mmol) and sodium-L-(+)-ascorbate (7 mg, 0.3 mmol) and
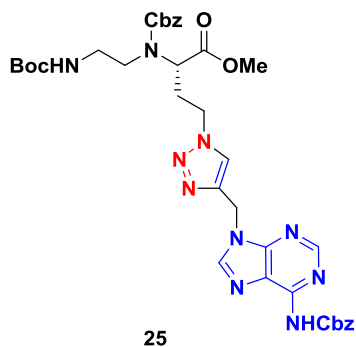
stirred. The resulting suspension was treated with propyne **12** (0.3 g, 0.6 mmol) dissolved in THF (5.0 ml) and stirred at room temperature. After 12 h, EtOAc (20 ml) was added into it and aq layer was further washed with EtOAc (3 x 20 ml). The organic layer was separated, dried over anhyd. Na_2SO_4, filtered and the filtrate was concentrated under reduced pressure. The crude mass thus obtained, was purified by column chromatography to afford

peptide **25** (0.4 g, 0.53 mmol 90%). White hygroscopic solid; IR (neat) 3391, 2967, 1744, 1700, 1614, 1517, 1463, 1363, 1213, 1159 1078, 987, 739, 699, 641, 589, 537 cm^{-1}; ^{1}H NMR (400 MHz, CDCl$_3$) δ 8.78 (s, 1H), 8.14 (s, 1H), 7.69 (d, J = 30.2 Hz, 1H), 7.46 – 7.28 (m, 10H), 5.48 (s, 2H), 5.29 (s, 2H), 5.19 – 5.01 (m, 3H), 4.40 (s, 2H), 4.06 (s, 1H), 3.79 – 3.57 (m, 3H), 3.56 – 3.45 (m, 1H), 3.22 (d, J = 30.7 Hz, 3H), 2.43 (ddd, J = 49.5, 49.1, 9.8 Hz, 3H), 1.39 (s, 9H); ^{13}C NMR (101 MHz, CDCl$_3$) δ 170.9, 156.0, 152.9, 151.2, 151.0, 149.5, 142.8, 142.1, 136.0, 135.5, 128.7, 128.6, 128.4, 128.0, 123.5, 121.9, 79.6, 68.0, 67.9, 66.0, 58.7, 52.9, 48.3, 47.5, 38.7, 28.5 ppm; HRMS (ESI-TOF) m/z calcd for $C_{36}H_{43}N_{10}O_8$ [M + H]$^+$ 743.3265, found 743.3260.

4-Triazolyl-[4-methyl-(N^6-benzyloxycarbonyl)-9H-purin-9-yl]-2(S)-[amino(N-ethylamino-t-butoxycarbonyl-N-benzyloxycarbonyl)] butanoic acid (9)
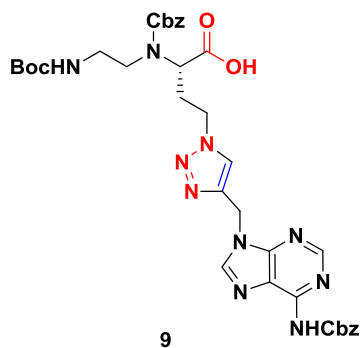
Compound **9** (0.26 g, 0.36 mmol) was prepared, purified, and characterized the same way as described for **4**, starting from compound **25** (0.30 g, 0.4 mmol), to afford yield (90%) IR (neat) 3401, 2923, 2857, 1682, 1614, 1535, 1461, 1399, 1209, 1147, 1076, 985, 841, 738, 697, 640. cm^{-1}; ^{1}H NMR (400 MHz, DMSO) δ 10.70 (s, 1H), 8.62 (d, J = 4.0 Hz, 1H), 8.51 (d, J = 6.2 Hz, 1H), 8.11 (d, J = 16.8 Hz, 1H), 7.50 – 7.21 (m, 10H), 7.12 (s, 1H), 5.57 (t, J = 35.2 Hz, 2H), 5.24 (d, J = 24.3 Hz, 2H), 4.98 (dd, J = 43.5, 13.1 Hz, 2H), 4.44 (m, 2H), 4.08 – 3.90 (m, 1H), 3.17 – 2.99 (m, 4H), 2.11 (s, 2H), 1.27 (t, J = 14.4 Hz, 9H); ^{13}C NMR (101 MHz, DMSO) δ 171.9 (s), 155.6 (s), 152.2 (s), 151.6 (s), 149.4 (s), 144.2 (s), 142.0 (s), 136.7, 136.4, 128.4, 128.3, 128.0, 127.9, 127.7, 127.2, 127.1, 123.8, 115.7, 77.8, 66.5, 66.3, 58.1, 53.6, 47.1, 28.2 ppm; HRMS (ESI-TOF) m/z calcd for $C_{35}H_{41}N_{10}O_8$ [M + H]$^+$ 729.3109, found 729.3110.

4-Triazolyl-[4-methyl(N⁶-benzyloxycarbonyl)-9H-purin-9-yl]-2(S)-[amino(N-ethylamino-t-butoxycarbonyl-N-acetamido(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)] methyl butanoate (26)
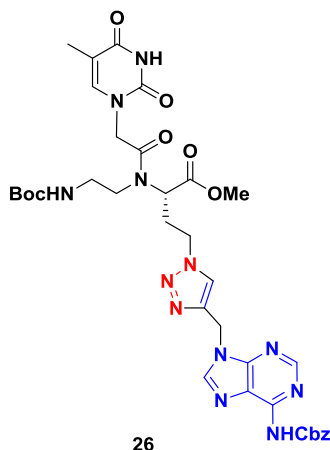
Compound **26** (0.64 g, 0.82 mmol) was prepared, purified, and characterized the same way as described for **25**, starting from compound **21** (0.43 g, 0.92 mmol), to afford yield (90%) IR (neat) 3854, 3737, 2931, 2312, 167, 1613, 1516, 1461, 1367, 1322, 1213, 1161, 1087, 991, 896, 736, 640, 549 cm^{-1}; ^{1}H NMR (400 MHz, CDCl$_3$) δ 10.23 (s, 1H), 9.07 (s, 1H), 8.77 (s, 1H), 8.29 (s, 1H), 7.75 (s, 1H), 7.51 – 7.29 (m, 5H), 6.88 (s, 1H), 5.58 – 5.44 (m, 3H), 5.27 (d, J = 12.4 Hz, 2H), 4.64 (d, J = 16.2 Hz, 1H), 4.49 – 4.27 (m, 2H), 4.10 (dd, J = 11.6, 8.7 Hz, 1H), 3.89 (s, 1H), 3.71 (s, 3H), 3.55 (dd, J = 13.5, 8.2 Hz, 1H), 3.49 – 3.30 (m, 1H), 3.10 (d, J = 9.3 Hz, 2H), 2.74 – 2.52 (m, 2H), 1.87 (s, 3H), 1.44 (s, 9H); ^{13}C NMR (101 MHz, CDCl$_3$) δ 170.6, 167.8, 164.8, 156.0, 153.0, 151.7, 151.4, 151.3, 149.7, 143.5, 142.7, 141.1, 135.7, 128.7, 128.5 , 128.4, 123.7, 121.6, 111.1, 80.2, 67.7, 59.5, 53.1, 49.1, 47.7, 39.1, 29.80, 29.5, 28.6, 22.8, 12.46, ppm; HRMS (ESI-TOF) m/z calcd for C$_{35}$H$_{43}$N$_{12}$O$_9$ [M + H]$^+$ 775.3276, found 775.3275.

4-Triazolyl-[4-methyl(N⁶-benzyloxycarbonyl)-9H-purin-9-yl]-2(S)-[amino(N-ethylamino-t-butoxycarbonyl-N-acetamido(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)] butanoic acid (10)

Compound **10** (0.29 g, 0.38 mmol) was prepared, purified, and characterized the same way as described for **4**, starting from compound **26** (0.35 g, 0.45 mmol), to afford yield (85 %) IR (neat) 3860, 3736, 3389, 2928, 2857, 2313, 1671, 1618, 1522, 1464, 1367, 1324, 1218, 1161, 1087, 1007, 901, 762, 640, 566 cm^{-1}; ^{1}H NMR (400 MHz, CDCl$_3$) δ 11.27 (d, J = 18.7 Hz, 1H), 10.66 (s, 1H), 8.62 (d, J = 5.9 Hz, 1H), 8.52 (d, J = 5.4 Hz, 1H), 8.18 (d, J = 33.4 Hz, 1H), 7.38 (m, 6H), 7.00 (d, J = 2.6 Hz, 1H), 5.54 (s, 2H), 5.21 (s, 2H), 4.61 (s, 1H), 4.56 – 4.30 (m, 3H), 4.01 (s, 1H), 3.46 – 2.94 (m, 8H), 1.75 (s, 3H), 1.33 (d, J = 24.0 Hz, 9H); ^{13}C NMR (101 MHz, DMSO) δ 167.3, 164.4, 155.7, 155.6, 152.1, 151.7, 151.6, 151.1, 151., 149.5, 144.2, 142.2, 142.1, 136.4, 128.4, 127.9, 127.8,

123.8, 123.7, 123.1, 108.1, 78.1, 77.6, 66.2, 48.0, 47.1, 38.3, 29.0, 28.1, 11.9, ppm; HRMS (ESI-TOF) m/z calcd for $C_{34}H_{41}N_{12}O_9$ [M + H]$^+$ 761.3119, found 761.3113.

*N*1(Boc)-1, 2-diaminoethane (28)

28
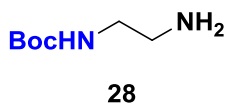
To an ice-cold stirred solution of 1,2-diamonoethane **27** (2.0 g, 33.3 mmol) in DCM (300 ml) was added solution of Boc-anhydride (5 g, 35 mmol) in DCM (50 ml) solution. The mixture was stirred for 12 h and the resulting solution was concentrated to100 ml. The *N*1,*N*2-di-Boc derivative not being soluble in H$_2$O, precipitated out and was removed by filtration. The corresponding *N*-mono-Boc derivative was obtained by repeated extraction from the filtrate in dichloromethane. Removal of solvents yielded the mono-Boc-diaminoethane **28** (3.4 g, 60%) which was used for further reaction without any purification.

Ethyl N2 (*t*-Boc-aminoethyl) glycinate (29)

29
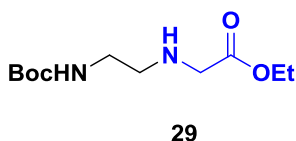
The *N1*(Boc)-1,2-diaminoethane **28** (3.2 g, 20 mmol) was treated with ethyl bromoacetate (2.2 ml, 20 mmol) in ACN (100 mL) in the presence of triethylamine (5.5 ml, 40 mmol) and the mixture was stirred at ambient temperature for 10 h. The completion of reaction was monitored by TLC. After completion of reaction, solvent was removed on a rotary evaporator. Reaction mixture was then diluted with H$_2$O (50 ml) and extracted with EtOAc (3 × 50 ml), followed by washing with brine. The collected organic layer was dried over anhydrous sodium sulphate and filtered.and purified with flash colum to afford (3.44 g, 13.97 mmol) compound **29** (yield 70 %), IR (neat) 3887, 3735, 3388, 2973, 2100, 1697, 1510, 1456, 1362, 1246, 1162, 1084, 1046, 1002, 864, 735, 699, 607 cm^{-1}; ^{1}H NMR (400 MHz, CDCl$_3$) δ 5.1 (s, 1H), 4.2 (m, 2H), 3.4 (s, 2H), 3.21 (d, *J* = 4 Hz, 2H), 2.74 (t, 2H) 1.44 (d, *J* = 31.0 Hz, 9H), 1.28 (t, 3H); ^{13}C NMR (101 MHz, CDCl$_3$) δ 173, 156, 79.4, 61.0, 50.6, 48.9, 40.3, 28.6, 14.4 ppm; HRMS (ESI-TOF) m/z calcd for $C_{34}H_{41}N_{12}O_9$ [M + H]$^+$ 761.3119, found 761.3113.

Ethyl [N-ethylamino(t-butoxycarbonyl), N(benzyloxy)carbonyl) glycinate (30)

30
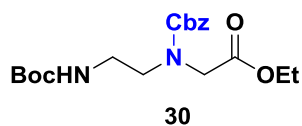
To a solution of **29** (0.8 g, 3.25 mmol) in DCM (20 ml) at 0 °C was added benzyl chloroformate (0.8 g, 5.0 mmol) dropwise followed by addition of N,N-diisopropylethylamine (1.7 g, 13.4 mmol). After

stirring at room temperature for 5 h, saturated aq. NH_4Cl solution (10 ml) was added and extracted with DCM (4 x 20 ml). The organic layers were dried over $MgSO_4$, concentrated *in vacuo* and purified with flash chromatography in hexane to afford **30** (Yield, 1.16 g, 3.0 mmol 92 %), TLC in EtOAc 4:1, R_f = 0.26; HRMS (ESI-TOF) m/z calcd for $C_{19}H_{28}N_2NaO_6$ $[M + Na]^+$ 403.1845, found 403.1844.

N-ethylamino(t-butoxycarbonyl) N(benzyloxy)carbonyl glycine (31)
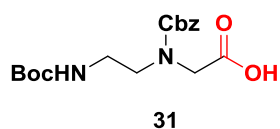
Compound **30** (0.87 g, 2.47 mmol) was prepared, purified, and characterized the same way as describe for compound **4**, starting from compound **30** (1.0 g, 2.63 mmol), to afford yield (94%); IR (neat) 3368, 2977, 1690, 151, 1464, 1413, 1364, 1239, 1157, 1075, 985, 857, 734, 696, 612 cm^{-1}; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.34 (t, *J* = 13.9 Hz, 5H), 6.49 (s, 1H), 5.14 (d, *J* = 12.9 Hz, 2H), 4.01 (d, *J* = 14.3 Hz, 2H), 3.47 (s, 2H), 3.24 (d, *J* = 23.7 Hz, 2H), 1.44 (d, *J* = 31.0 Hz, 9H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 173.3, 156.3, 136.3, 128.7, 128.6, 128.1, 127.8, 79.8, 67.8, 48.8, 40.8, 39.3, 28.5, ppm; HRMS (ESI-TOF) m/z calcd for $C_{17}H_{24}N_2NaO_6$ $[M + Na]^+$ 375.1532, found 375.1532.

Ethyl cytosine-1-yl acetate (33)

To a suspension of cytosine **32** (5.0 g, 45.0 mmol) in dry DMF (100 ml) was added NaH (1.08 g, 45.0 mmol) under N_2 at 0 °C. The mix was stirred for 2 h at rt then ethyl bromo acetate (4.3 ml, 45.0 mmol) was added. The mix was stirred for 48 h at rt. The solvent was then evaporated in vacuo. The crude residue was triturated with H_2O (100 ml) and the precipitate was filtered off. Recrystallization from MeOH/H_2O gave of compound **33** (5.27 g, 26.7 mmol 60%) as white crystals (mp: 225 – 227 °C); HRMS (ESI-TOF) m/z calcd for $C_8H_{12}N_3O_3$ $[M + H]^+$ 198.0878, found 198.0887.

Ethyl (N^4 –benzyloxycarbonyl) cytidin-1-yl acetate (34)
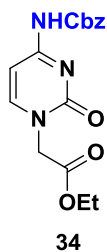
Cbz-Cl (3.12 ml, 21.8 mmol) and DMAP (2.67 g, 21.8 mmol) were dissolved at -15 °C in DCM (20 ml). The mixture was stirred for 15 min and compound **33** (2.0 g, 10.15 mmol) was gradually added. After stirring for 15 min at -15 °C then 5 h at rt, the mixture was evaporated in vacuo and the crude residue was taken up in $CHCl_3$. The organic layer was washed with 1 M HCl, with H_2O then dried over

Na$_2$SO$_4$ and lastly evaporated in reduced pressure. Trituration of the crude residue in ether gave a white precipitate which was then filtered off, affording **34** (2.59 g, 7.82 mmol 77%) as an amorphous powder. TLC (EtOAc/MeOH 1:1) R$_f$ = 0.79. ^{1}H NMR (400 MHz, DMSO) δ 10.82 (s, 1H), 8.04 (d, *J* = 7.3 Hz, 1H), 7.45 – 7.33 (m, 5H), 7.04 (d, *J* = 7.2 Hz, 1H), 5.19 (s, 2H), 4.60 (s, 2H), 4.14 (q, *J* = 7.1 Hz, 2H), 1.20 (t, *J* = 7.1 Hz, 3H); ^{13}C NMR (101 MHz, DMSO-D6) δ 168.0, 163.5, 155.1, 153.2, 150.5, 136.0, 128.6, 128.3, 128.0, 94.2, 66.6, 61.2, 50.7, 14.1 ppm.

N^4(Benzyloxycarbonyl) cytidin-1-yl acetic acid (35)
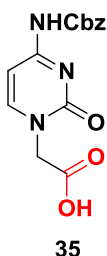
Compound **34** (2.0 g, 6.0 mmol) was dissolved in dioxane (50 ml) and aqueous 1 M NaOH (8.8 ml) was added. The mix was stirred 5 h at rt and then concentrated under reduced pressure. The residue was taken up in a aqueous KHSO$_4$ (1 M). The resultant white precipitate was isolated by filtration, washed with H$_2$O and dried in a vacuum desiccator affording **35** (1.59 g, 5.24 mmol, 87%). TLC (EtOAc/MeOH 1:1) R$_f$ = 0.10. ^{1}H NMR (400 MHz, DMSO) δ 10.84 (s, 1H), 8.04 (d, *J* = 7.3 Hz, 1H), 7.46 – 7.31 (m, 5H), 7.02 (d, *J* = 7.3 Hz, 1H), 5.20 (s, 2H), 4.53 (s, 2H); ^{13}C NMR (101 MHz, DMSO-D6) δ 169.4, 163.4, 155.1, 153.2, 150.5, 136.0, 128.5, 128.2, 128.0, 94.1, 66.6, 50.6 ppm; HRMS (ESI-TOF) m/z calcd for C$_{14}$H$_{14}$N$_3$O$_5$ [M + H]$^+$ 304.0933, found 304.0934.

N^6(Benzyloxycarbonyl)adenine (37)
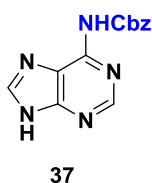
Sodium hydride (6.1 g, 0.25 mol, 60% dispersion in oil) was washed with petroleum ether (3 x). After cooling in an ice bath, dry DMF (150 ml) was added, followed by adenine **36** (5.0 g, 37 mmol) in small portions. The suspension was stirred vigorously for 3 min and then benzyl chloroformate (11.6 mL, 81.5 mmol) was added dropwise. After stirring for 4 h the reaction mixture was poured into ice H$_2$O (300 mL) and the pH adjusted to 7 with aq. HCl (1 N). The light yellow precipitate was collected by filtration and washed with H$_2$O, and ether to afford crude product (8.5 g). Recrystallization from MeOH/CHCl$_3$ afforded two crops of **37** (4.2 g, 16.0 mmol 44%) as a white solid: m.p = 222 °C; ^{1}H NMR (400 MHz, DMSO) δ 12.25 (s, 1H), 11.05 (s, 1H), 8.59 (s, 1H), 8.44 (s, 1H), 7.61 – 7.31 (m, 5H), 5.29 (s, 2H); ^{13}C NMR (101 MHz, DMSO-D6) δ 162.3, 154.2, 151.7, 146.5, 145.0, 136.4, 129.0, 128.8, 128.6, 113.3, 67.4 ppm; HRMS (ESI-TOF) m/z calcd for C$_{13}$H$_{12}$N$_5$O$_2$ [M + H]$^+$ 270.0991, found 270.0981.

N[6](Benzyloxycarbonyl)-N9-(methyl-t-butylcarboxy) adenine (38)
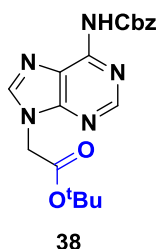
To a suspension of **37** (0.5 g, 1.9 mmol) in dry DMF (4 ml) was added anhydrous Cs_2CO_3 (0.06 g, 0.19 mmol) and anhydrous K_2CO_3 (0.3 g, 1.9 mmol). After stirring for 5 min, tert-butyl bromoacetate (0.3 ml, 2.1 mmol) was added dropwise. After stirring for ca.23 h, the mixture was evaporated to dryness in vacuo and the residue partitioned between EtOAc (35 ml) and H_2O (10 ml). The organic phase was washed with H_2O and brine, dried (Na_2SO_4) and concentrated in vacuo. The residue was crystallized from EtOAc/hexane to afford **38** (0.4 g, 1.0 mmol 55%) as a white solid: m.p = 141-142 °C; ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.77 (s, 1H), 8.58 (s, 1H), 7.98 (s, 1H), 7.47 – 7.33 (m, 5H), 5.30 (s, 2H), 4.86 (s, 2H), 1.47 (s, 9H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 165.9, 153.1, 151.7, 151.0, 149.5, 143.4, 135.6, 128.8, 128.8, 128.7, 121.5, 84.0, 67.9, 45.0, 28.1 ppm; HRMS (ESI-TOF) m/z calcd for $C_{19}H_{22}N_5O_4$ [M + H]$^+$ 384.1672, found 384.1670.

N[6](Benzyloxycarbonyl)-N9-(carboxymethyl) adenine (39)
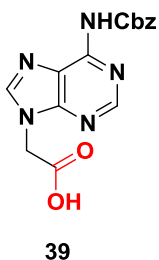
A solution of **38** (0.15 g, 0.39 mmol) in dry DCM (3.1 ml) was treated with triethylsilane (0.6 ml), cooled to 0 °C and TFA (1.5 ml) was added slowly. After 5 min the reaction was allowed to warm to rt. After 7.5 h, the mixture was evaporated to dryness in vacuo and any remaining volatiles removed by azeotroping with $CHCl_3$ (3 x) to give **39** (0.12 g, 0.38 mmol, 97%) as a white foam. An analytically pure sample of **39** was obtained by recrystallization from acetone: m.p = 139-143 °C; ^{1}H NMR (400 MHz, DMSO) δ 13.39 (s, 1H), 10.72 (s, 1H), 8.62 (s, 1H), 8.44 (s, 1H), 7.49 – 7.29 (m, 5H), 5.22 (s, 2H), 5.09 (s, 2H); ^{13}C NMR (101 MHz, DMSO) δ 169.0, 152.2, 151.6, 149.4, 144.8, 136.3, 128.4, 128.2, 128.0, 127.8, 122.9, 66.3, 44.3 ppm; HRMS (ESI-TOF) m/z calcd for $C_{15}H_{14}N_5O_4$ [M + H]$^+$ 328.1046, found 328.1046.

N[2](Isobutyryl) guanine (41)

Isobutyric acid anhydride (21.3 g, 0.13 mol) was added at room temperature to a suspension of guanine **40** (7.6 g, 50.3 mmol) in N, N-dimethylacetamide (100 ml). The reaction mixture was heated at 150 °C for 2 hours, cooled to room temperature and evaporated under reduced pressure to 1/10 of its volume. The precipitated crude product was collected and crystallised from boiling

EtOH:H_2O (1:1, 150 ml) to yield **41** (10. g, 45.2 mmol, 89%) as colorless crystals, and used for next step, without any purification.

N^2(Isobutyryl)-N9-(methyl-t-butylcarboxy) guanine (42)
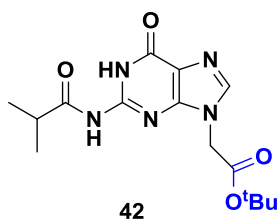
To a suspension of N^2-isobutyrylguanine **41** (6.6 g, 30.0 mmol) in dry DMF (250 ml) was added NaH (0.7 g, 30.0 mmol) at 0 °C. The mixture was stirred for 1 h and then tert-butyl bromoacetate (6.4 g, 33.0 mmol) was added dropwise over a period of 30 min at 0 °C. The reaction was stopped after 2 h by the addition of a small amount of solid CO_2 and MeOH (5 ml). The reaction mixture was evaporated in vacuo and the residue was treated with DCM and H_2O. The H_2O phase was re-extracted with DCM. The DCM fractions were combined and concentrated in vacuo to give the crude product. The resulting crude product was recrystallized from EtOAc to give the desired N9-isomer product **42** with a yield (6.5 g 19.39 mmol 64%) mp = 204–205 °C (decomp.); R_f = 0.13 (DCM/MeOH, 9.5:0.5); ^{1}H NMR (400 MHz, DMSO) δ 12.10 (s, 1H), 11.65 (s, 1H), 7.95 (s, 1H), 4.88 (s, 2H), 2.86 – 2.69 (m, 1H), 1.42 (s, 9H), 1.11 (d, J = 6.8 Hz, 6H); ^{13}C NMR (101 MHz, DMSO-d_6) δ 180.2(s), 166.7, 154.8, 148.9, 148.1, 140.3, 119.6, 82.3, 44.8, 34.7), 27.7, 18.9 ppm; HRMS (ESI-TOF) m/z calcd for $C_{15}H_{22}N_5O_4$ $[M + H]^+$ 336.1672, found 336.1671.

N^2-(isobutyryl)-N9-(carboxymethyl) guanine (43)
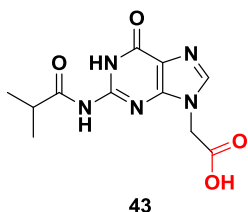
To a suspension of **42** (2.0 g, 6.0 mmol) dry DCM (10 ml) was added triethylsilane (4.8 ml, 30 mmol). The mixture was cooled to 0 °C and TFA (15 ml) was then added over a period of 5 min. After 30 min the reaction was allowed to warm to room temperature. Upon completion of the reaction (traced by TLC), the mixture was concentrated in vacuo to give a foam which was stirred with diethyl ether, filtered, washed with diethyl ether, and dried in vacuo to afford the desired product **43** with a yield (1.67 g, 6.0 mmol, 100%). R_f = 0.44 (n-butanol/acetic acid/H_2O, 3:1:1); ^{1}H NMR (400 MHz, DMSO) δ 12.08 (s, 1H), 11.67 (s, 1H), 7.95 (s, 1H), 4.90 (s, 2H), 2.86 - 2.69 (m, 1H), 1.11 (d, J = 6.8 Hz, 6H); ^{13}C NMR (101 MHz, DMSO-d_6) δ 180.2, 169.1, 154.8, 149.0, 148.1, 140.3, 119.7, 44.4, 34.7, 18.9 ppm; HRMS (ESI-TOF) m/z calcd for $C_{11}H_{14}N_5O_4$ $[M + H]^+$ 280.1046, found 280.1049.

Thymin-1-yl acetic Acid (45)
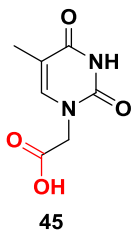
To a suspension of thymine **44** (5.0 g, 39.7 mmol) and K_2CO_3 (5.5 g, 40 mmol) in dry DMF (120 ml) was added ethyl bromoacetate (37.5 ml, 0.39 mol), and the mixture was stirred vigorously overnight under N_2. The mixture was filtered and evaporated to dryness, in vacuo. The solid residue was cooled to 0 °C, treated with H_2O (37 ml) and aq. HC1 (4 N, 1.6 ml), and stirred for 30 min. The precipitate was collected by filtration and washed with H_2O (3 x 20 ml). The precipitate was treated with H_2O (40 ml) and aq. NaOH (2 M, 20 ml) and boiled for 10 min. The mixture was cooled to 0 °C, treated with aq. HCl (4 M, 13 ml), and stirred for 30 min. The compound **45** was collected by filtration, washed with H_2O (3 x 20 ml), and dried to afford (4.5 g, 24.45 mmol, 62%); ^{1}H NMR (400 MHz, DMSO-d_6) δ 13.11 (s, 1H), 11.33 (s, 1H), 7.49 (d, J = 1.1 Hz, 1H), 4.36 (s, 2H), 1.75 (d, J = 0.9 Hz, 3H).HRMS (ESI-TOF) m/z calcd for $C_7H_9N_2O_4$ [M + H]$^+$ 185.0562, found 185.0566.

N1(t-butoxycarbonyl)-N4,N4-di-[*t*-butoxycarbonyl)-*2-oxopyrimidine-1(2H)-carboxylate* (47)

To a 100 mL Ar-flushed flask equipped with a magnetic stir bar and containing cytosine **46** (1.0 g, 9.0 mmol) and DMAP (0.1 g, 0.9 mmol) was added dropwise dry THF (50 ml) to the stirred suspension was added Boc_2O (7.85 g, 36 mmol) under an Ar atmosphere. The reaction mixture was stirred overnight at room temperature (TLC analysis indicated the presence of a single product and starting solid cytosine was entirely solubilized). The excess amount of THF was evaporated, and the crude product was dissolved in EtOAc (300 ml), washed with HCl aq. (1 N, 1 × 15 ml) and brine (2 × 15 ml), dried with Na_2SO_4 and concentrated in vacuo to give **47** as a pale yellow oil (3.67 g, 8.9 mmol, 99%). R_f = 0.45 (hexane/EtOAc, 7:3). ^{1}H NMR (CDCl$_3$): δ = 7.97 (d, J = 7.9 Hz, 1 H), 7.07 (d, J = 7.8 Hz, 1 H), 1.60 (s, 9 H), 1.56 (s, 18 H) ppm. ^{13}C NMR (CDCl$_3$): δ = 162.3, 160.0, 150.1, 149.0, 143.3, 96.8, 85.2, 27.5 ppm. HRMS (ESI): calcd for $C_{19}H_{29}N_3O_7$ [M]$^+$ 411.2006; found 411.1907.

N^4-bis-(t-Boc) cytidine (48)

To a solution of **47** (8.0 g, 19.4 mmol) dissolved in MeOH (300 ml) was added saturated aq. $NaHCO_3$ (90 ml). The turbid solution was stirred at 50 °C for 1 h, at which point clean conversion to bis-Boc protected cytosine **48** was observed by

TLC. After evaporation of MeOH, H_2O (130 ml) was added to the suspension, and the aqueous layer was extracted with CH_2Cl_2 (3 x 100 ml). The organic layer was dried with Na_2SO_4, filtered and evaporated to give pure **48** (5.99 g, 19.25 mmol 99%) as a white solid. R_f = 0.65 (AcOEt). mp = 320 °C (decomp.), and further used for next step without any purification.

N^4-bis-(t-Boc)-N1(propyne-2-yl) cytidine (11)
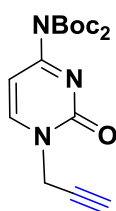
To a mixture of **48** (3.0 g, 9.6 mmol) and dry K_2CO_3 (1.4 g, 10.1 mmol) in dry DMF (25 ml) was added propargyl bromide (80% in toluene, 2.6 ml, 20 mmol). The resulting mixture was stirred at rt. for 18 h. The reaction mixture was suspended in H_2O (100 ml) and the aq. layer was extracted with EtOAc (50 x 3 ml). The combined organic layer was washed with H_2O and brine solution. The organic layer was separated, dried over Na_2SO_4, filtered and the filtrate was concentrated under reduced pressure. The crude mass thus obtained, was purified by column chromatography [Eluent: 20-50% of EtOAc in pet ether] to afford compound **11** (3.15 g, 9.0 mmol, 93%). White solid; m.p. 147 °C – 150 °C; ^{1}H NMR, 400 MHz (DMSO-d_6, 25°C): δ = 8.24 (d, J = 7.3 Hz, 1H), 6.85 (d, J = 7.3 Hz, 1H), 4.66 (d, J = 2.5 Hz, 2H), 3.49 (t, J = 2.5 Hz, 1H), 1.49 ppm (s, 18H); ^{13}C NMR, 100 MHz (DMSO-d_6, 25°C): δ = 161.9, 153.4, 149.6, 149.2, 96.0, 84.6, 78.0, 76.6, 38.8, 27.2 ppm. HRMS (ESI$^+$), m/z calculated for (M+Na)$^+$ $C_{17}H_{23}N_3O_5Na$: 372.1535, found: 372.1542.

N^6-(benzyloxycarbonyl)-N9-(propyne-2-yl) adenine (12)
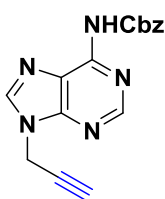
To a suspension of N^6(Benzyloxycarbonyl)adenine **37** (0.4 g, 1.5 mmol) in dry DMF (4 ml) was added dry K_2CO_3 (0.2 g, 1.5 mmol). After stirring for 5 min, propargyl bromide 80% in toluene (0.17 ml, 1.5 mmol) was added dropwise. After stirring for 5 - 7 h, the mixture was evaporated to dryness in vacuo and the residue partitioned between EtOAc (5 mL) and H_2O (2 mL). The organic phase was washed with H_2O and brine, dried (Na_2SO_4) and concentrated in vacuo. The residue was purified from EtOAc/hexane (1:1) to afford **12** (0.26 g, 0.85mmol 56%) as a yellowish oily compound. IR (neat) 3392, 2922, 2856, 1674, 1621, 1463, 1395, 1215, 1157, 1076, 740, 696, 637, 592, 539 ^{1}H NMR (400 MHz, CDCl$_3$) δ 9.50 (s, 1H), 8.79 (s, 1H), 8.12 (s, 1H), 7.42 – 7.33 (m, 5H), 5.29 (s, 2H), 4.93 (d, J = 2.6 Hz, 2H), 2.53 (t, J = 2.6 Hz, 1H). ^{13}C NMR (101 MHz, CDCl$_3$) δ 153.1, 151.3, 151.1, 149.8, 142.2, 135.5, 128.7 , 128.7 , 128.6, 121.9,

76.5 , 75.5, 67.8, 33.3. HRMS (ESI-TOF) m/z calcd for $C_{16}H_{14}N_5O_2$ [M + H]$^+$ 308.1147, found 308.1154.

2-Amino-6-chloro-N9(propyne-2-yl) purine (50)
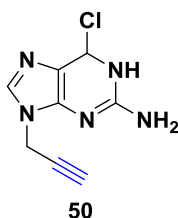
2-Amino-6-chloropurine **49** (3.0 g, 17.7mmol) was suspended in dry DMF (50 mL) followed by addition of anhydrous K_2CO_3 (2.9 g, 1.2 eq) and stirring under N_2 atmosphere for 1 hr, after this propargyl bromide (1.9 g, 0.9 eq) is added and stirred for 48 hours under N_2 atmosphere at room temperature. After then DMF was evaporated at 60 °C under high vacuum and compound was purified by column chromatography eluting with MeOH/CHCl$_3$ to afford yellowish-white solid **50** (2.4 g, 11.5 mmol, 65%) mp. 160-165°C; ^{1}H NMR (400 MHz, DMSO-d$_6$, 25°C): δ = 8.17 (s, 1H), 7.02 (s, 2H), 4.92 (d, *J* = 2.5 Hz, 2H), 3.48 ppm (t, *J* = 2.5 Hz, 1H); ^{13}C NMR (100 MHz, DMSO-d$_6$, 25°C): δ = 159.9, 153.6, 149.5, 142.4, 123.1, 77.9, 76.1, 32.4 ppm; HRMS (ESI$^+$), m/z calculated for (M+H)$^+$ $C_8H_7ClN_5$: 208.0390, found: 208.0397.

N9(propyne-2-nyl) guanine (13)
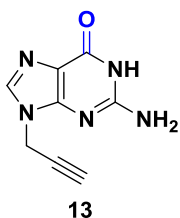
Compound **50** (1.0 g, 4.8 mmol) was dissolved in 3:1 mixture of TFA-H$_2$O (10 ml) and then stirred for 48 h at ambient temperature. The reaction mixture was evaporated and washed well with diethyl ether to afford compound **13** as brownish white solid, (0.75 g, 3.85 mmol, 80%) m.p. 167-170 °C; ^{1}H NMR (400 MHz, DMSO-d$_6$, 25 °C): δ = 11.45 (s, 1H), 8.69 (s, 1H), 7.16 (bs, 2H), 4.94 (d, *J* = 2.5 Hz, 2H), 3.60 ppm (t, *J* = 2.5 Hz, 1H); ^{13}C NMR (100 MHz, DMSO-d$_6$, 25°C): δ = 155.1, 154.5, 149.9, 136.7, 110.9, 77.1, 33.4 ppm; HRMS (ESI$^+$), m/z calculated for (M+H)$^+$ $C_8H_9N_5O$: 190.0729, found: 190.0735.

2.6 References

1. Kumar, V. A.; Ganesh, K. N. *Acc. Chem. Res.* **2005**, *38,* 404–412.

2. a) Haaima, G.; Lohse, A.; Buchardt, O.; Nielsen, P. E. *Angew. Chem., Int. Ed.* **1996**, *35*, 1939–1942. b) Puschl, A.; Sforza, S.; Haaima, G.; Dahl, O.; Nielsen, P. E. *Tetrahedron Lett*. **1998**, *39*, 4707-4710. c) Sforza, S.; Galaverna, G.; Dossena, A.; Corradini, R.; Marchelli, R. *Chirality* **2002**, *14*, 591–598.

3. Gourishankar, A.; Ganesh, K. N. *Artificial DNA: PNA & XNA* **2012**, *3*, 5–13.

4. Sugiyama, T.; Imamura, Y.; Demizu, Y.; Kurihara, M.; Takano, M.; Kittaka, A. *Bioorg*. *Med*. *Chem*. *Lett*. **2011**, *21*, 7317–7320.

5. Ganesh, K. N.; Nielsen, P. E. *Curr*. *Org*. *Chem*. **2000,** *4*, 931–943.

6. Sforza, S.; Galaverna, G.; Dossena, A.; Corradini, R.; Marchelli, R. *Chirality* **2002**, *14*, 591–598.

7. Hyrup, B.; Egholm, M.; Rolland, M.; Nielsen, P. E.; Berg, R. H.; Buchardt, O. *J. Chem.Soc. Chem. Commun.* **1993**, 518–519.

8. Hyrup, B.; Egholm, M.; Nielsen, P.E.; Wittung, P.; Norden, B.; Buchardt, O. *J. Am. Chem. Soc*. **1994,** *116*, 7964–7968.

9. (a) Haaima, G.; Lohse, A.; Buchardt, O.; Nielsen, P. E. *Angew*. *Chem*. *Int*. *Ed*. **1996**, *35*, 1939-1942. 10. (a) Mitra, R.; Ganesh, K. N. *Chem Commun*. **2011**, *47*, 1198–1200. (b) Sforza, S.; Tedeschi, T.; Corradini, R.; Marchelli, R. Eur. *J*. *Org*. *Chem*. **2007**, 5879–5885.

11. Gasser, G.; Husken, N.; Koster, S. D.; Metzler-Nolte, N. *Chem. Commun.* **2008**, 3675-3677.

12. Mitra, R.; Ganesh, K. N. *J*. *Org*. *Chem*. **2012**, *77*, 5696–5707.

13. Sahu, B.; Sacui, I.; Rapireddy, S.; Zanotti, K.; Bahal, J. R.; Armitage, B. A. Ly, D. H. *J*. *Org*. *Chem*. **2011**, *76*, 5614–5627.

14. Kramer, R. A.; Bleicher, K. H.; Wennemers, H. *Helv. Chim. Acta* **2012**, *95*, 2621-2634.

15. Porcheddu, A.; Giacomelli, G.; Piredda, I.; Carta, M.; Nieddu, G. *Eur. J. Org. Chem*. **2008**, 5786-5797.

16. Halay, E.; Ay, E.; Salva, E.; Ay, K.; Karayildirim, T. *Nucleosides Nucleotides Nucleic Acids*, **2017**, *36*, 598-619.

17. Nagapradeep, N.; Verma, Sandeep. *Chem. Commun.* **2011**, *47*, 1755-1757.

18. Zhang, L. H.; Kauffman, G. S.; Pesti, J. A.; Yin, J. *J. Org. Chem*. **1997**, *62*, 6918–6920.

19. Vilaivan, T. *Acc. Chem. Res.* **2015**, *48*, 1645-1656.

20. Bernardes, G. J. L.; Linderoth, L.; Doores, K. J.; Boutureira, O.; Davis, B. G. *ChemBioChem* **2011**, *12*, 1383–1386.

21. Mitra, R.; Ganesh, K. N. *Chem. Commun.* **2011**, *47*, 1198–1200.

22. Wang, Y.; Kirschner, A.; Fabian, A. K.; Gopalakrishnan, R.; Kress, C.; Hoogeland, B.; Koch, U.; Kozany, C.; Bracher, A.; Hausch, F. *Eur. J. med. Chem.* **2013**, *56*, 3922–3935.

23. Sahu, B.; Sacui, I.; Rapireddy, S.; Zanotti, K. J.; Bahal, R.; Armitage, B. A.; Ly, D. H. *J. Org. Chem.* **2011**, *76*, 5614–5627.

24. Prokhorov D. I.; Kirillova Y. G.; Boyarskaya N. P.; Tevyashova A. N.; Esipova O. V.; Zvonkova E. N.; Shvets V. I. *Pharm. Chem. J.* **2005**, *39*, 323–328.

25. Cooper, D. C.; Suggs, J. W. *Tetrahedron Lett.* **2012**, *53*, 6943-6945.

26. Narindoshvili, A. R. T. *Org. Biomol. Chem.* **2008**, *6*, 3171–3176.

27. Thomson, S. A.; Josey, J. A.; Cadilla, R.; Gaul, M. D.; Fred Hassman, C.; Luzzio, M. J.; Pipe, A. J.; Reed, K. L.; Ricca, D. J.; Wiethe, R. W.; Noble, S. A. *Tetrahedron* **1995**, *51*, 6179–6194.

28. Englund, E. A.; Xu, Q.; Witschi, M. A.; Appella, D. H. *J. Am. Chem. Soc.* **2006**, *128*, 16456–16457.

29. Hudson, R. H. E.; Goncharenko, M.; Wallman, A. P.; Wojciechowski, F. *Synlett* **2005**, *2005*, 1442–1446.

30. N. Nagapradeep, S. Verma, *Chem. Commun.* **2011**, *47*, 1755-1757;

31. R. H. E. Hudson, M. Goncharenko, A. P. Wallman, F. Wojciechowski, *Synlett* **2005**, *2005*, 1442-1446.

32. Will, D. W.; Breipohl, G.; Langner, D.; Knolle, J.; Uhlmann, E. *Tetrahedron* **1995**, *51*, 12069-12082

33. Jenny, T. F.; Schneider, K. C.; Benner, S. A. *Nucleosides and Nucleotides* **1992**, *11*, 1257-1261.

34. Porcheddu, A.; Giacomelli, G.; Piredda, I.; Carta, M.; Nieddu, G. *Eur. J. Org. Chem.* **2008**, *2008*, 5786–5797.

2.7 Appendix I: Characterization data of synthesized compounds

| Entry	Table of contents	Page No.
1	^{1}H, ^{13}C, IR and HRMS spectra of all compounds	87 - 132

¹H NMR of Compound 18
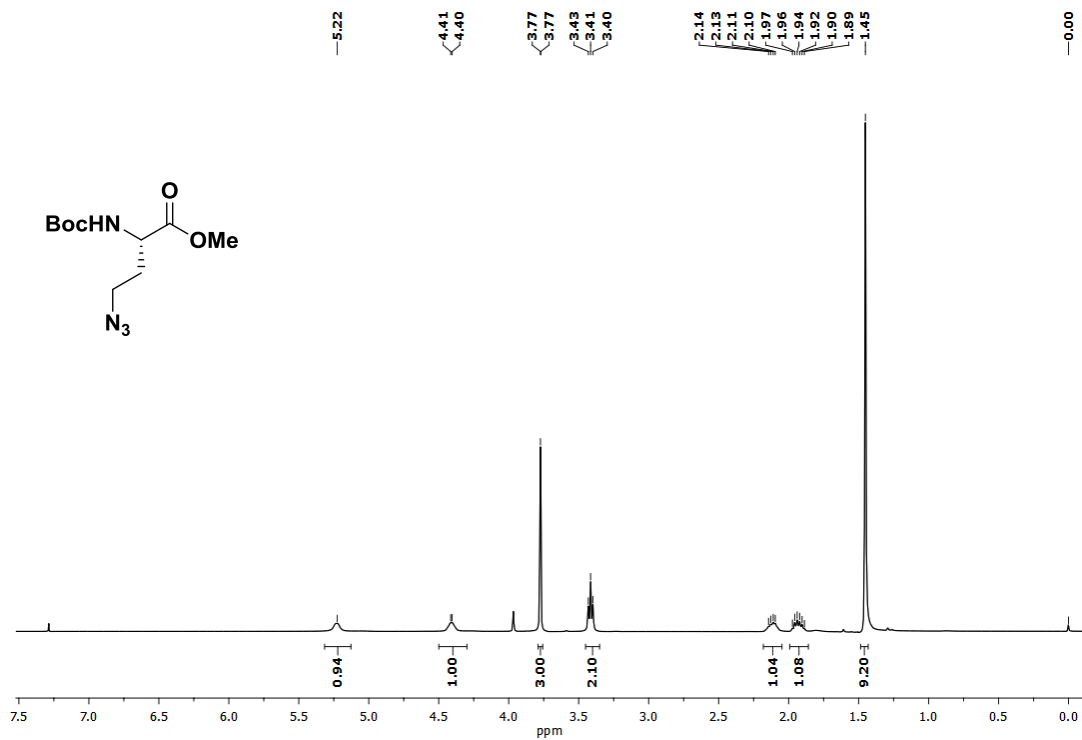
¹³C NMR of Compound 18
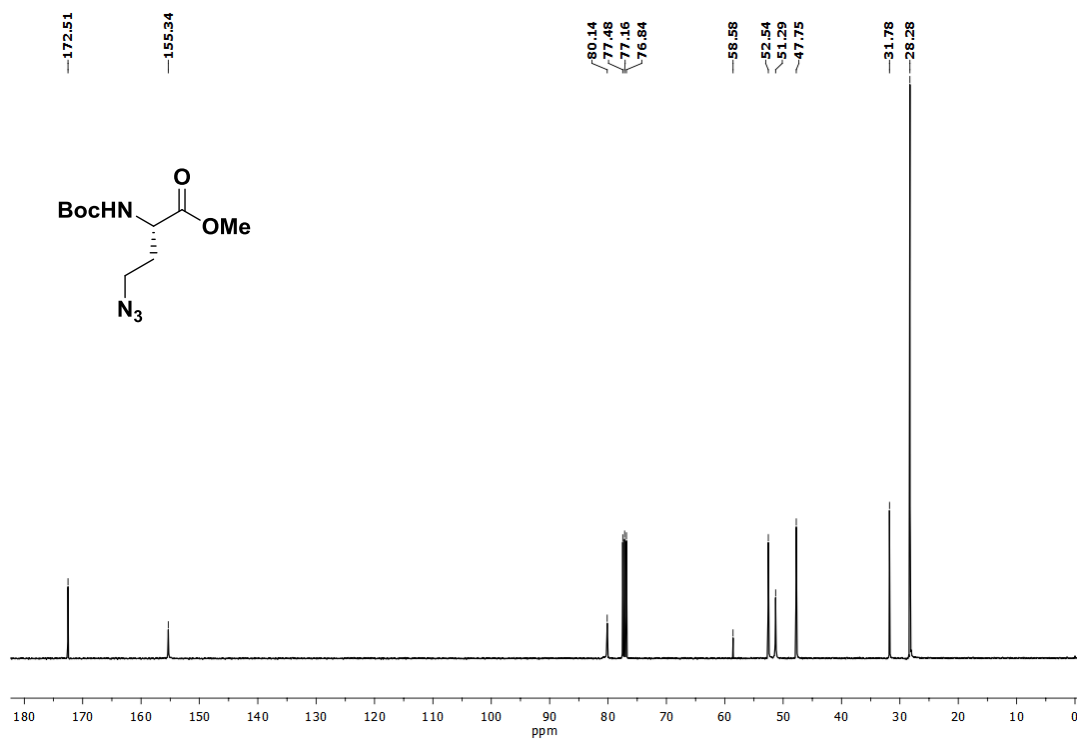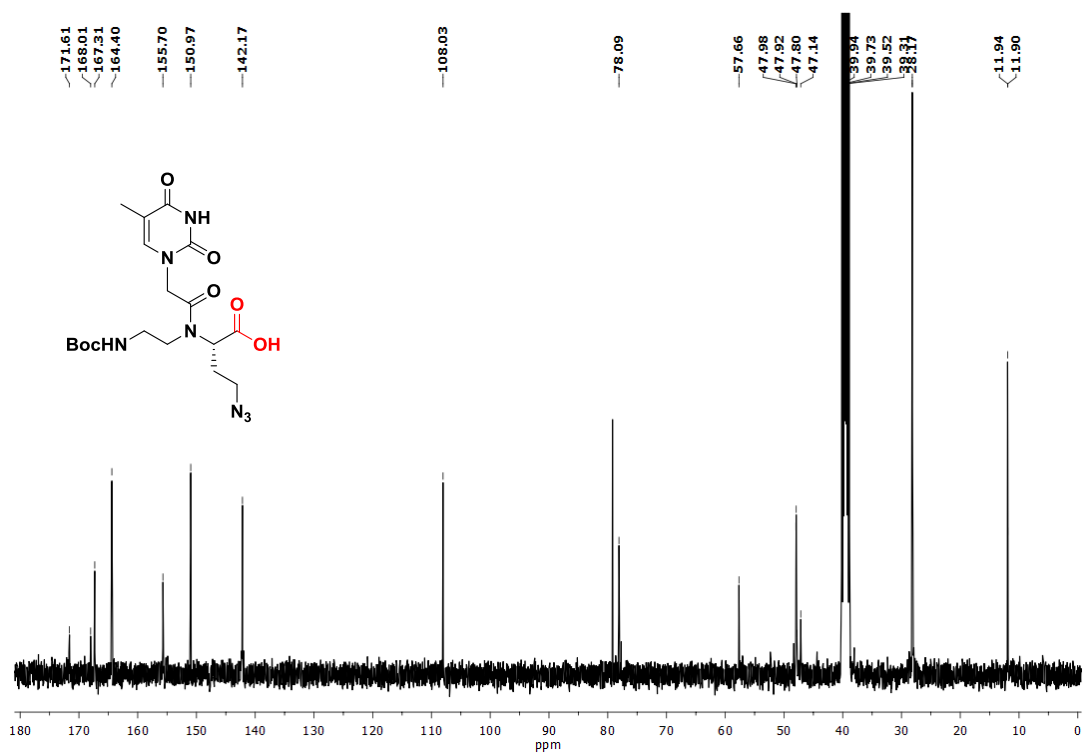
IR Spectrum of Compound 18

HRMS of Compound 18

Calculated mass $(M+Na)^+ = 281.1226$
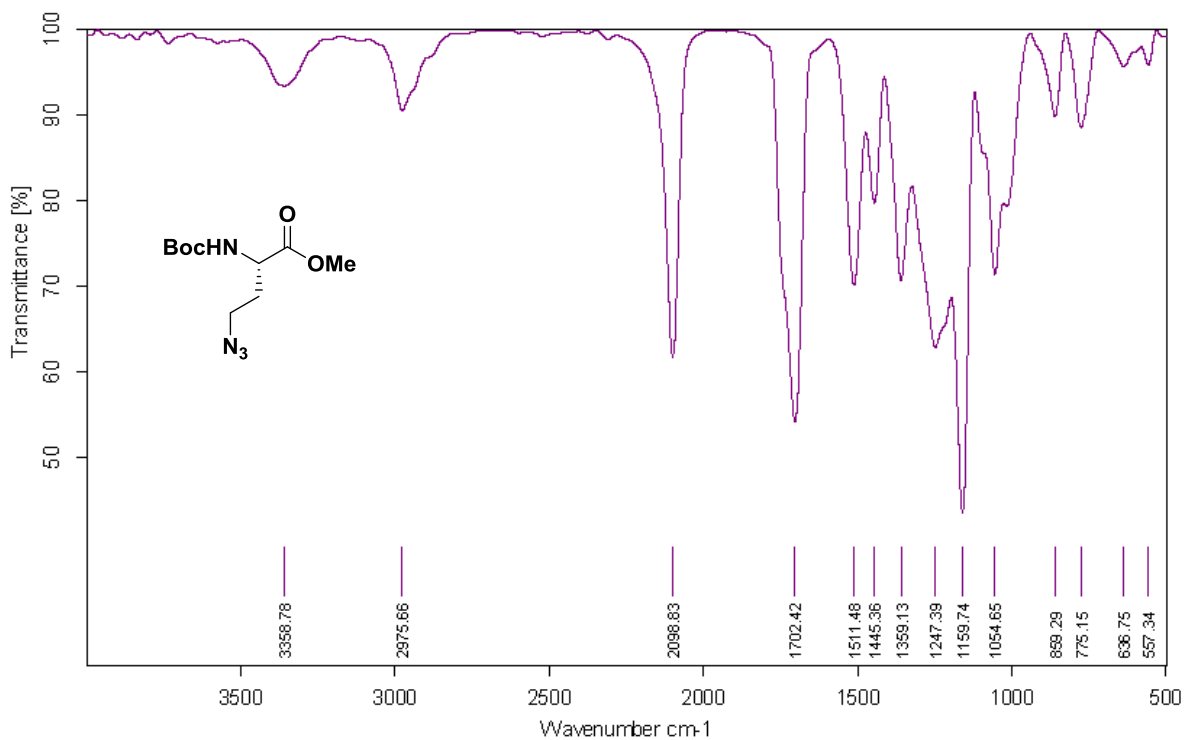
Obs. mass $(M+Na)^+ = 281.1218$
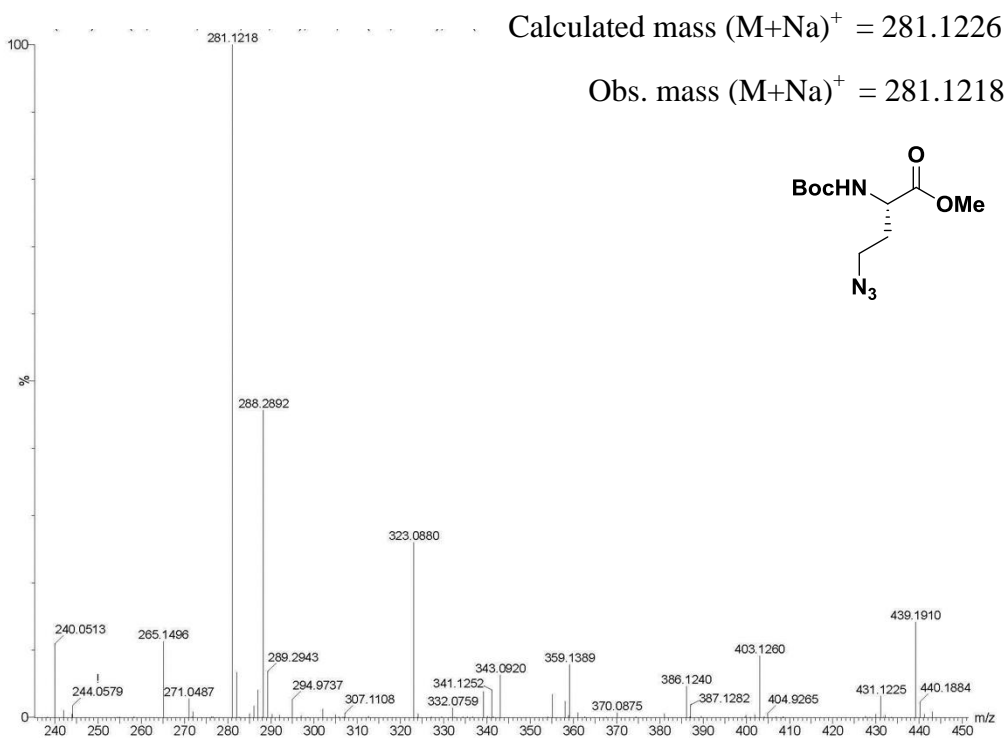
^{1}H NMR of Compound 19
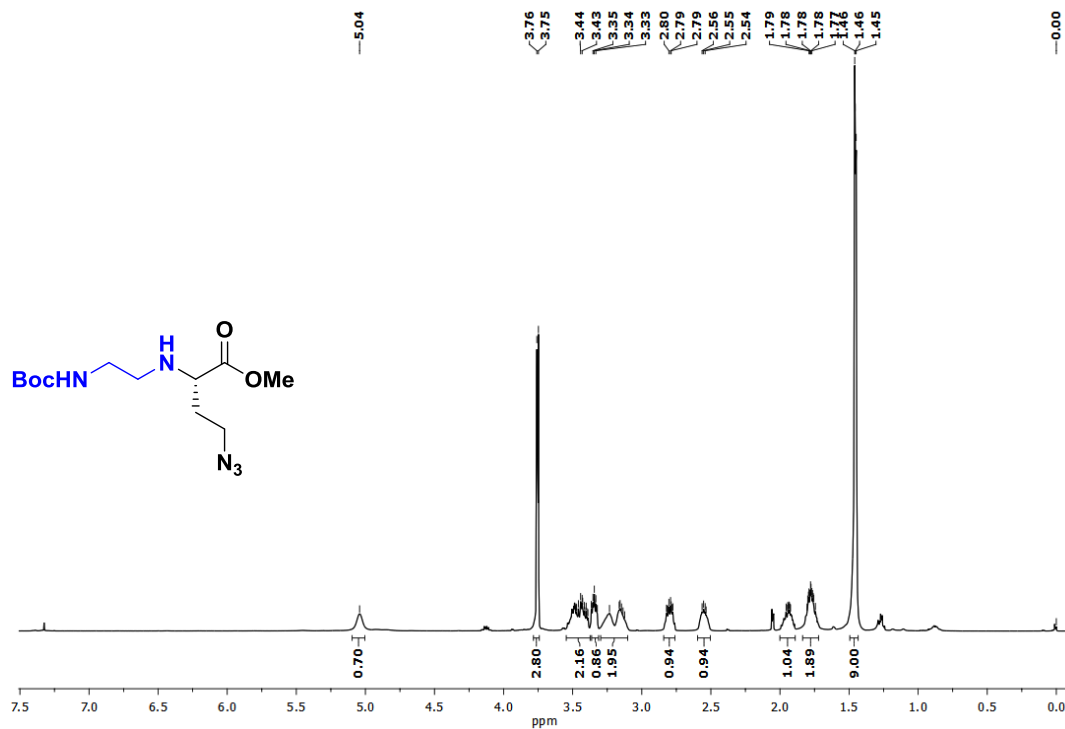
^{13}C NMR of Compound 19
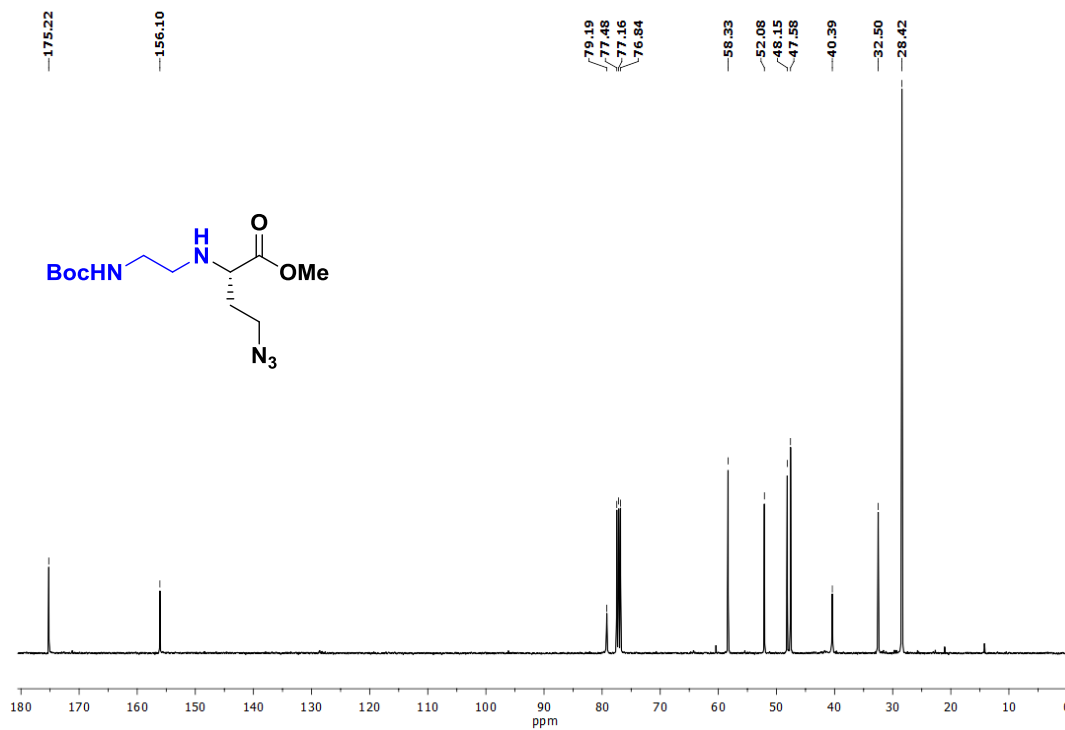
IR Spectrum of Compound 19

HRMS of Compound 19

Calculated mass $(M+H)^+ = 302.1828$

Obs. mass $(M+H)^+ = 302.1832$
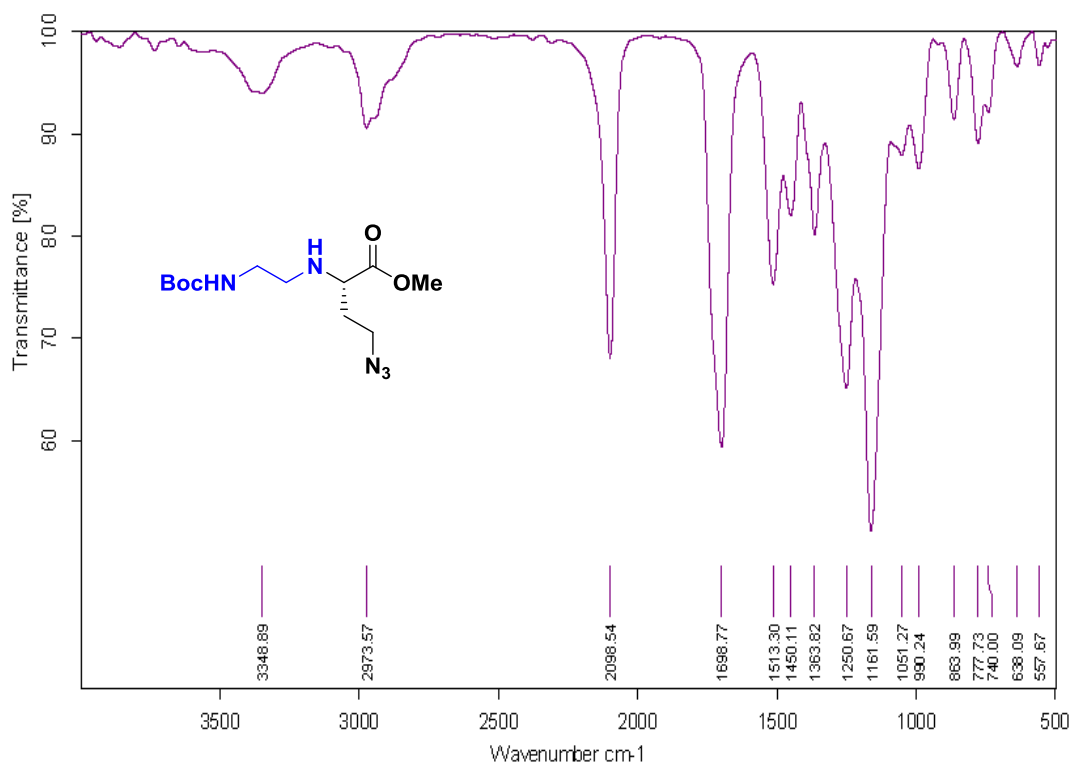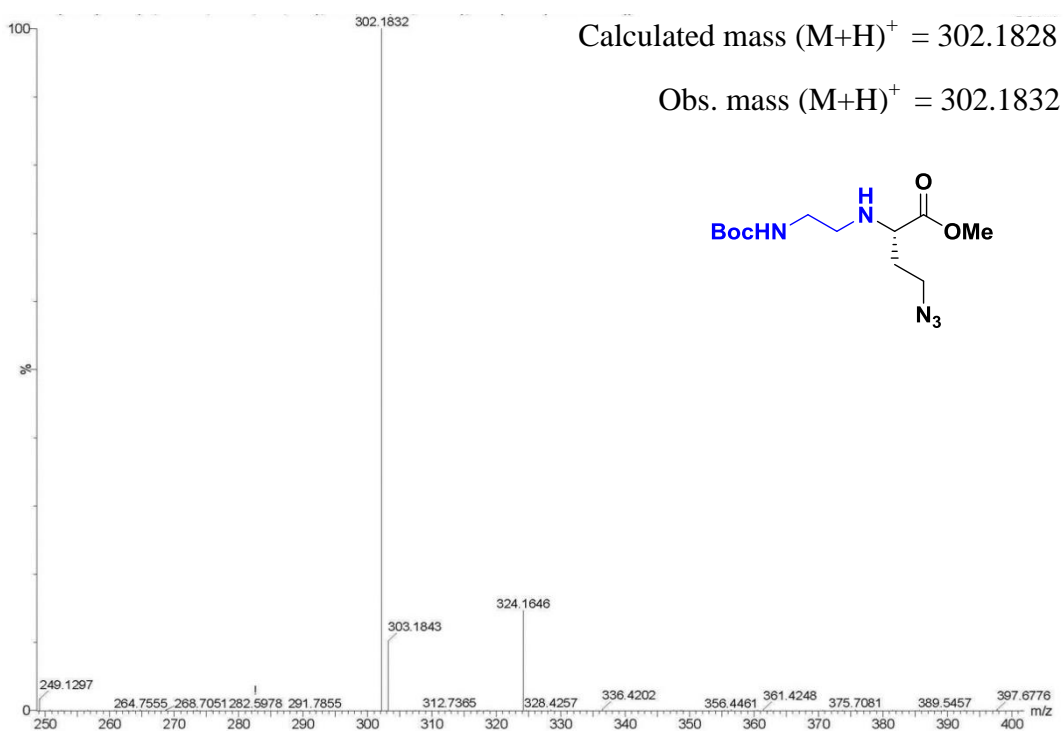
^{1}H NMR of Compound 20
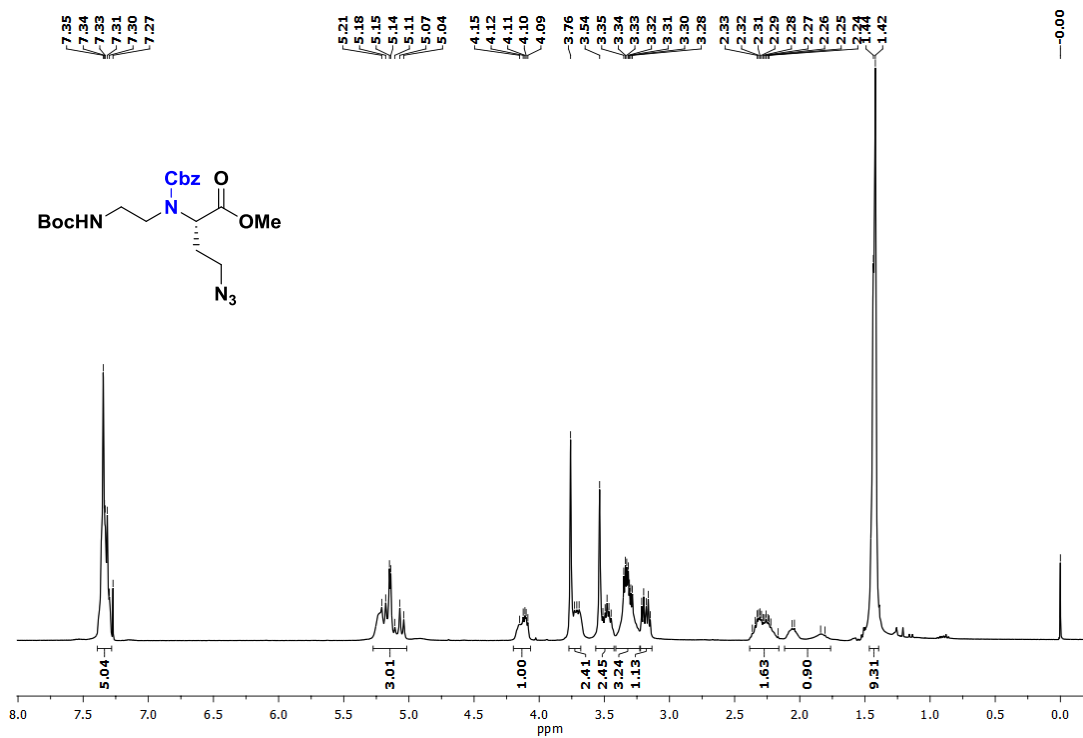
^{13}C NMR of Compound 20
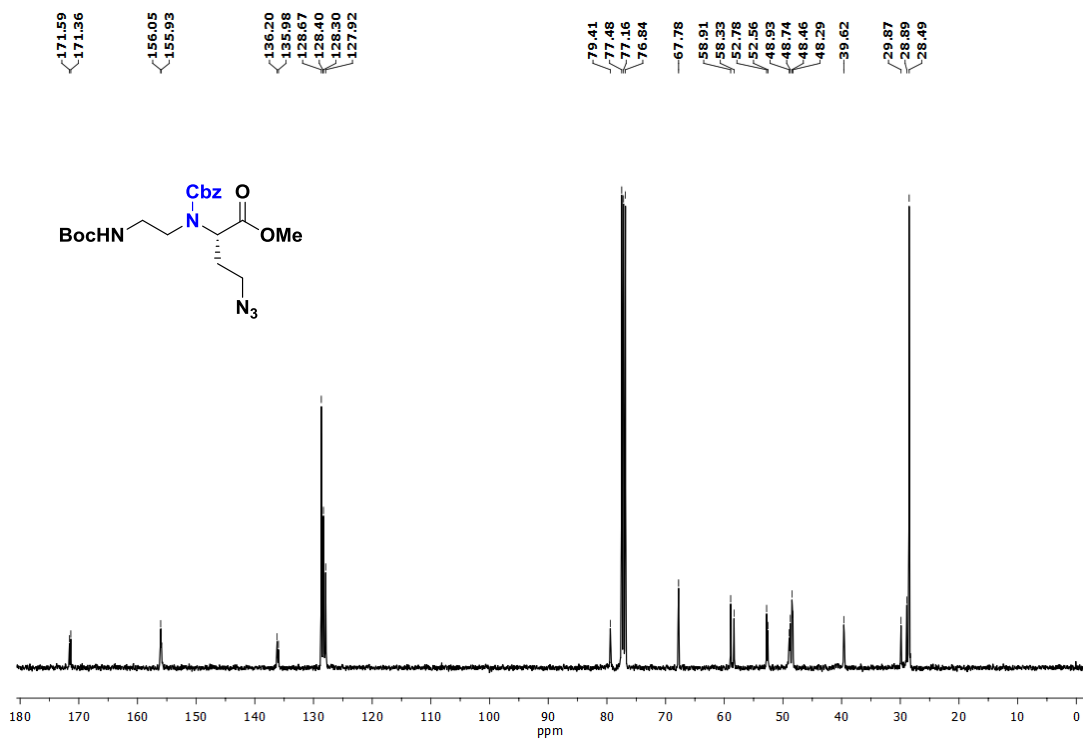
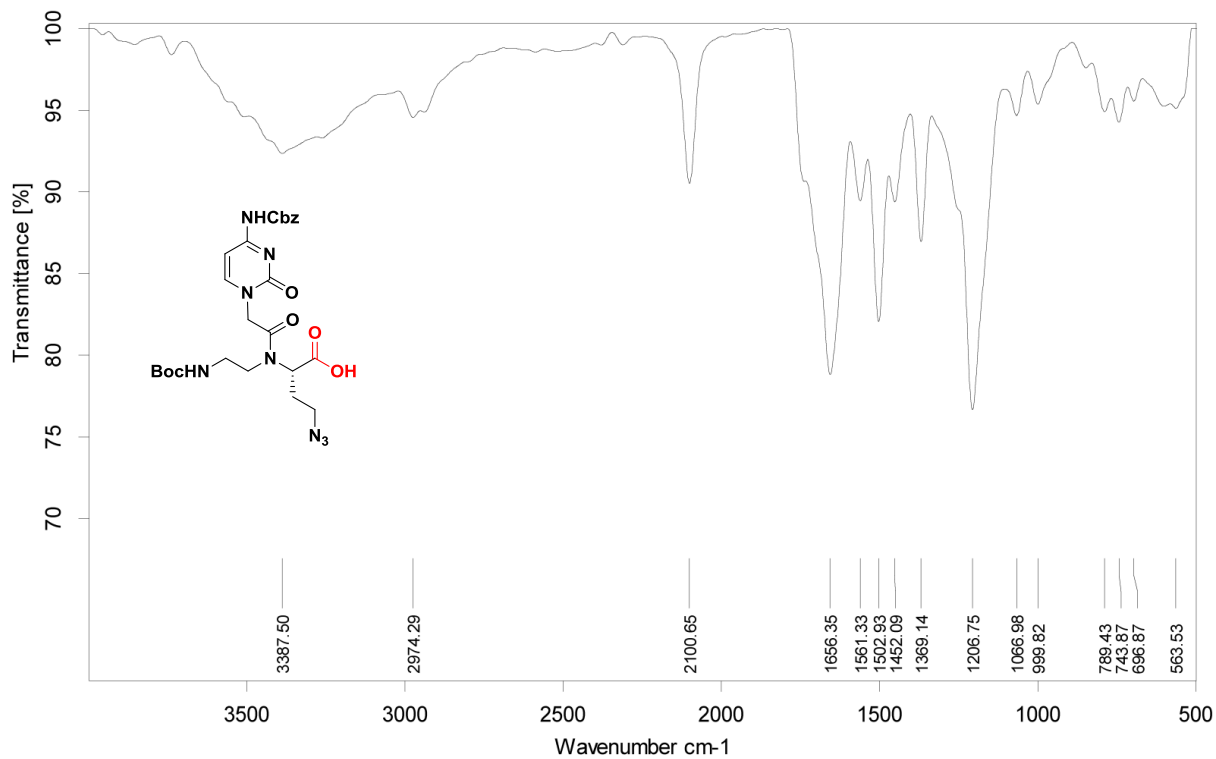## IR spectrum of Compound 20
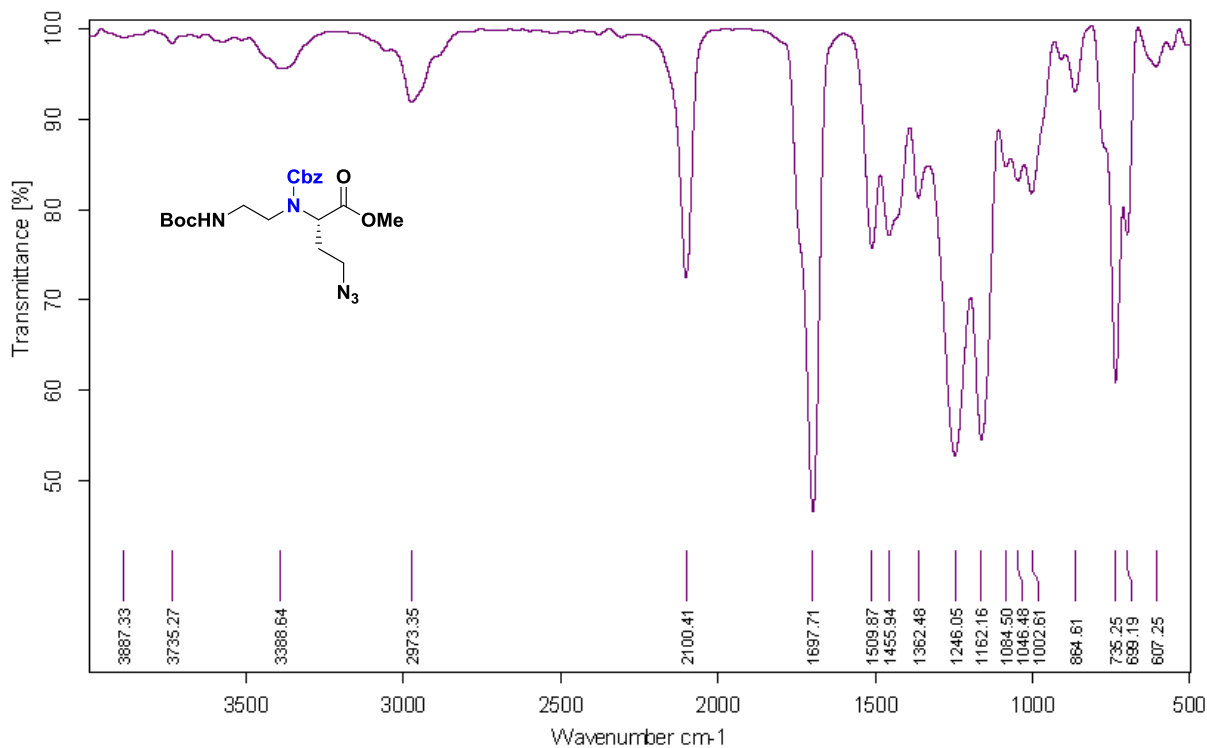
¹H NMR of Compound 8
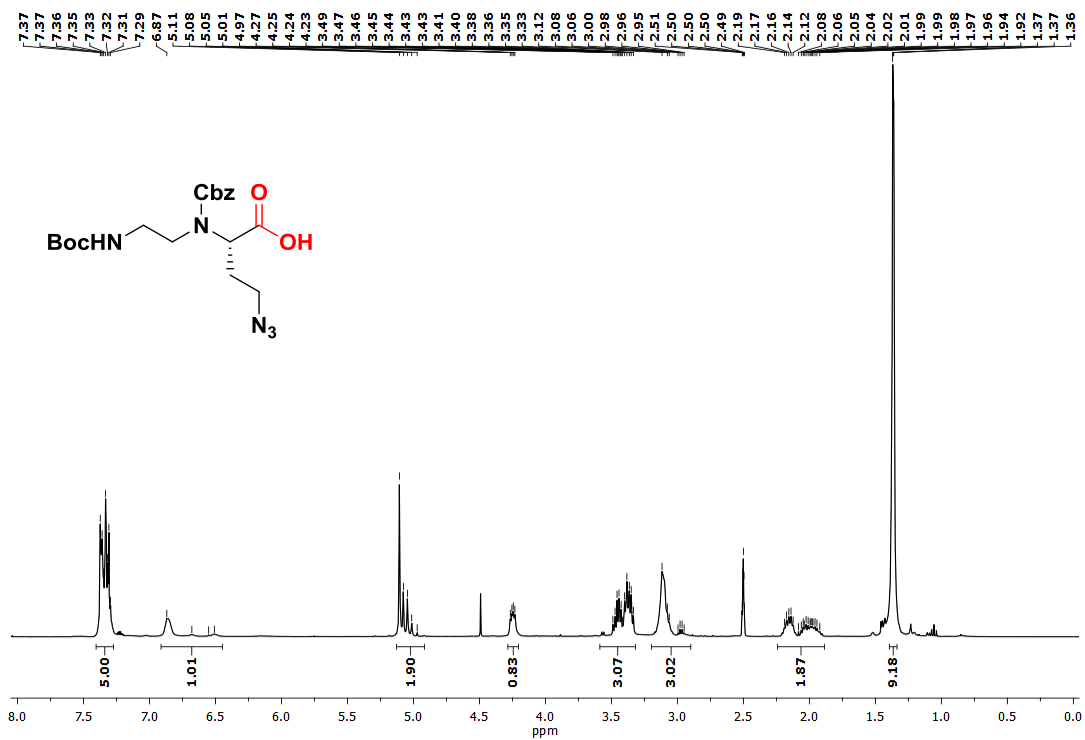
¹³C NMR of Compound 8
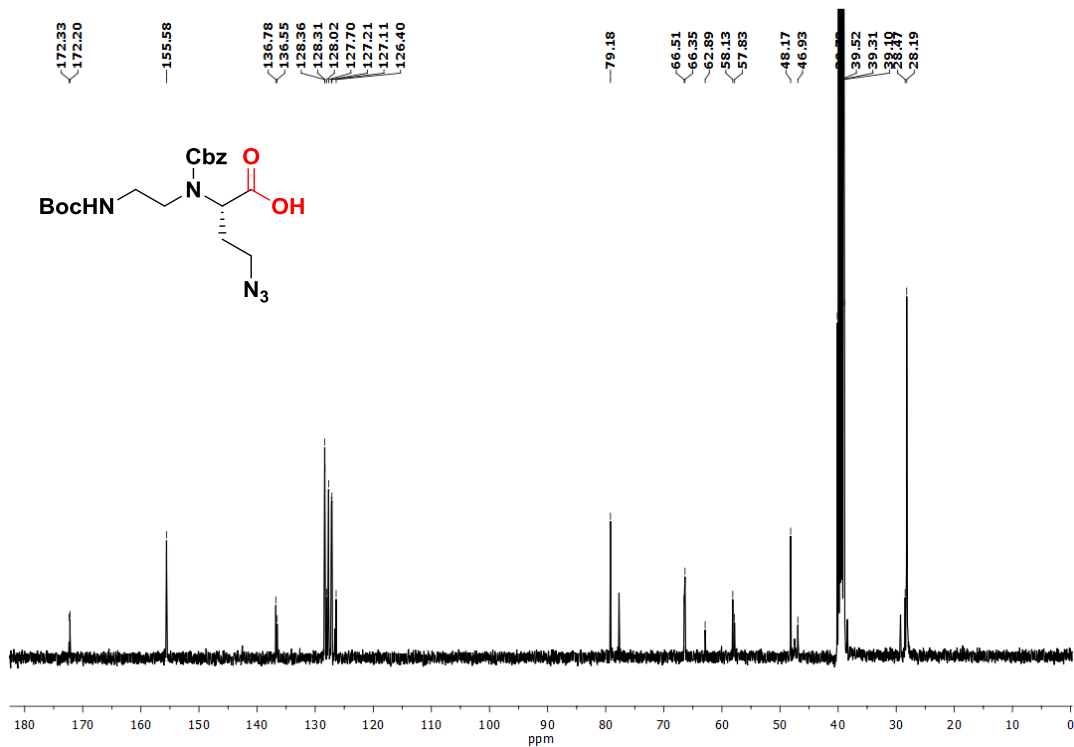
IR Spectrum of Compound 8
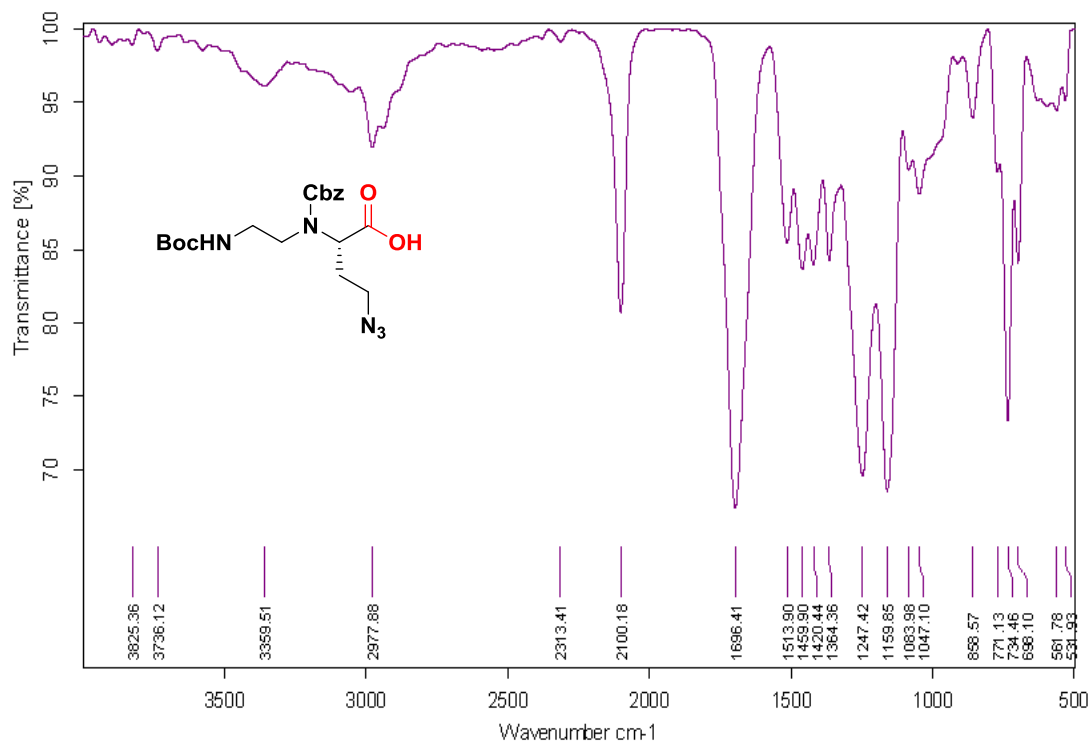
¹H NMR of Compound 21
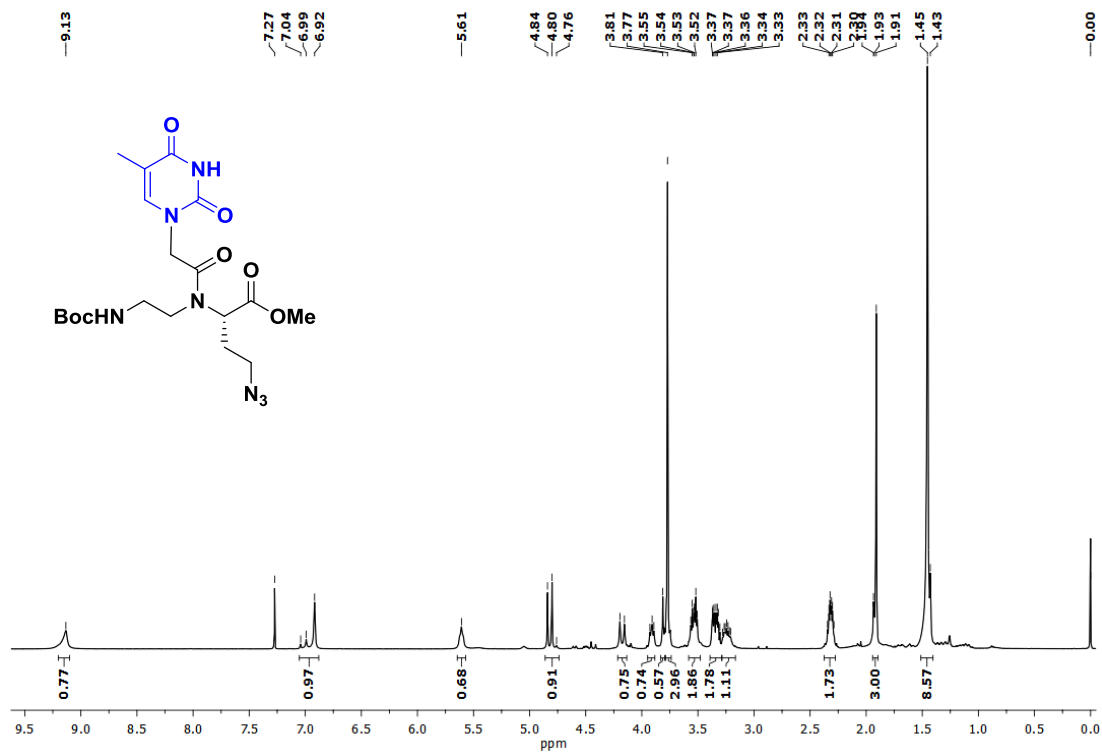
¹³C NMR of Compound 21
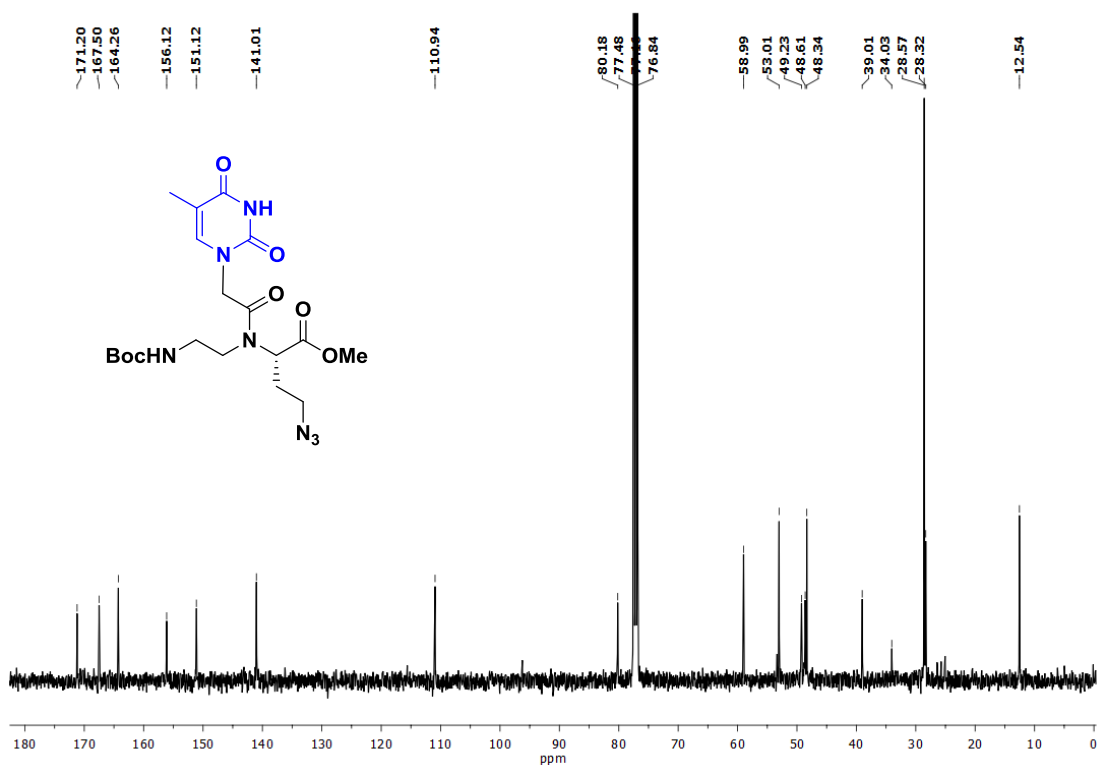
IR Spectrum of Compound 21

HRMS of Compound 21

Calculated mass $(M+Na)^+ = 490.2026$

Obs. mass $(M+Na)^+ = 490.2026$
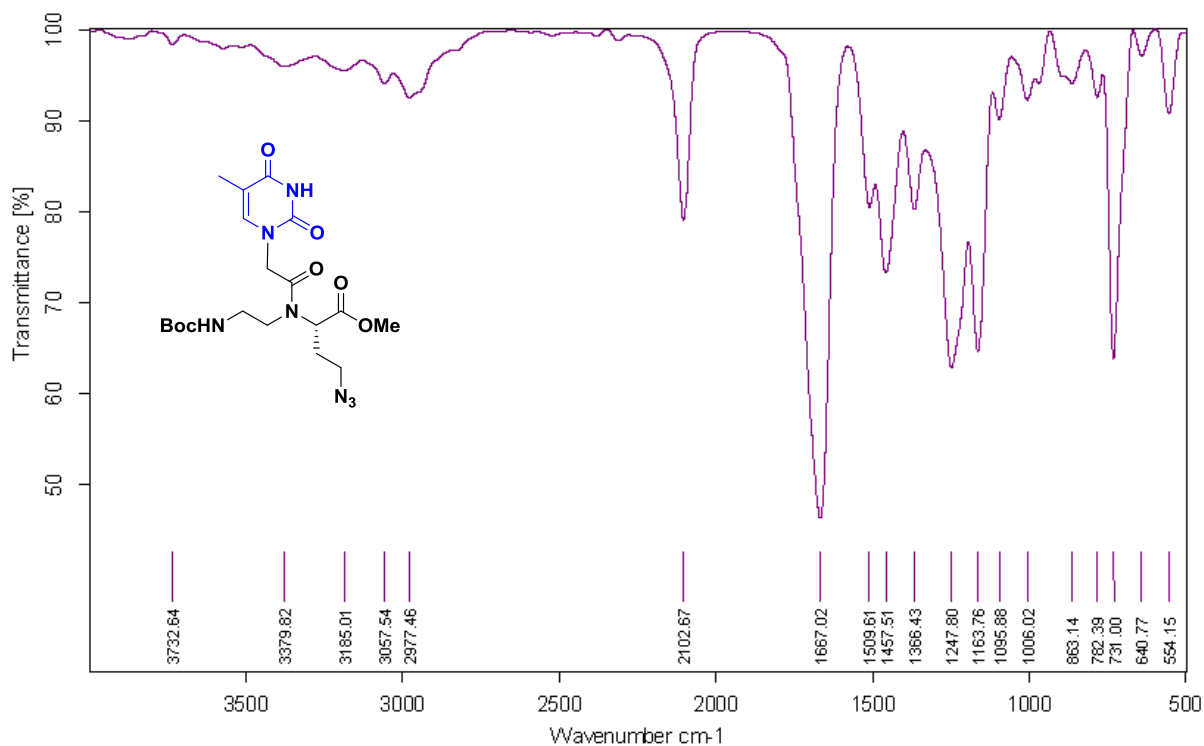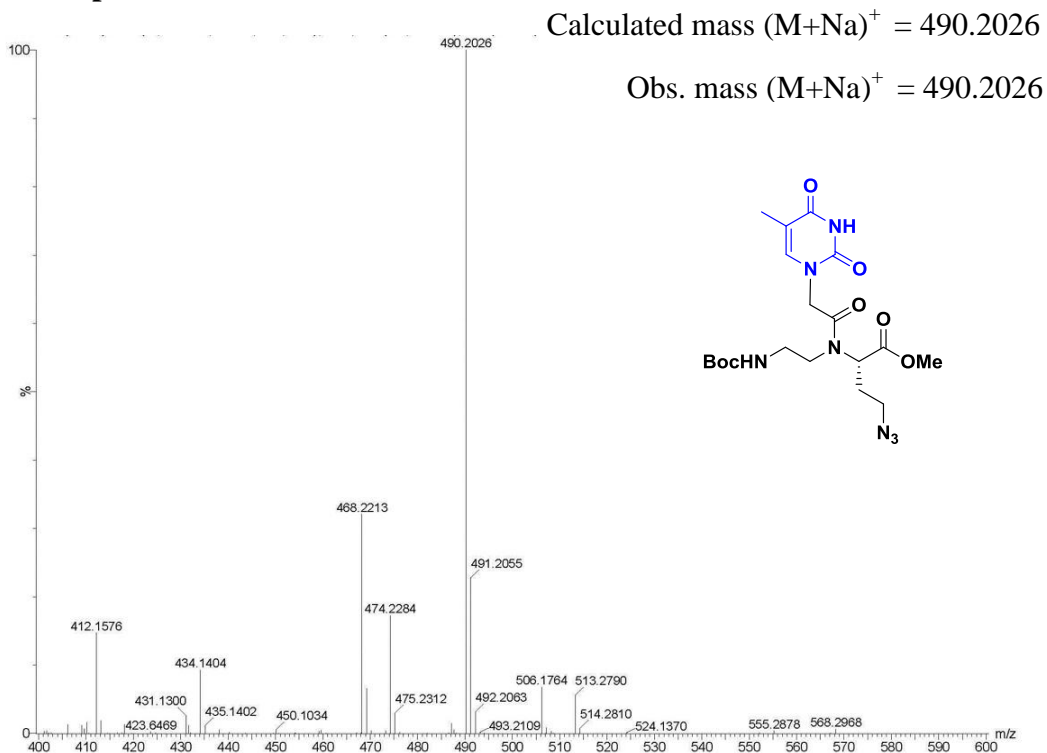
^{1}H NMR of Compound 22
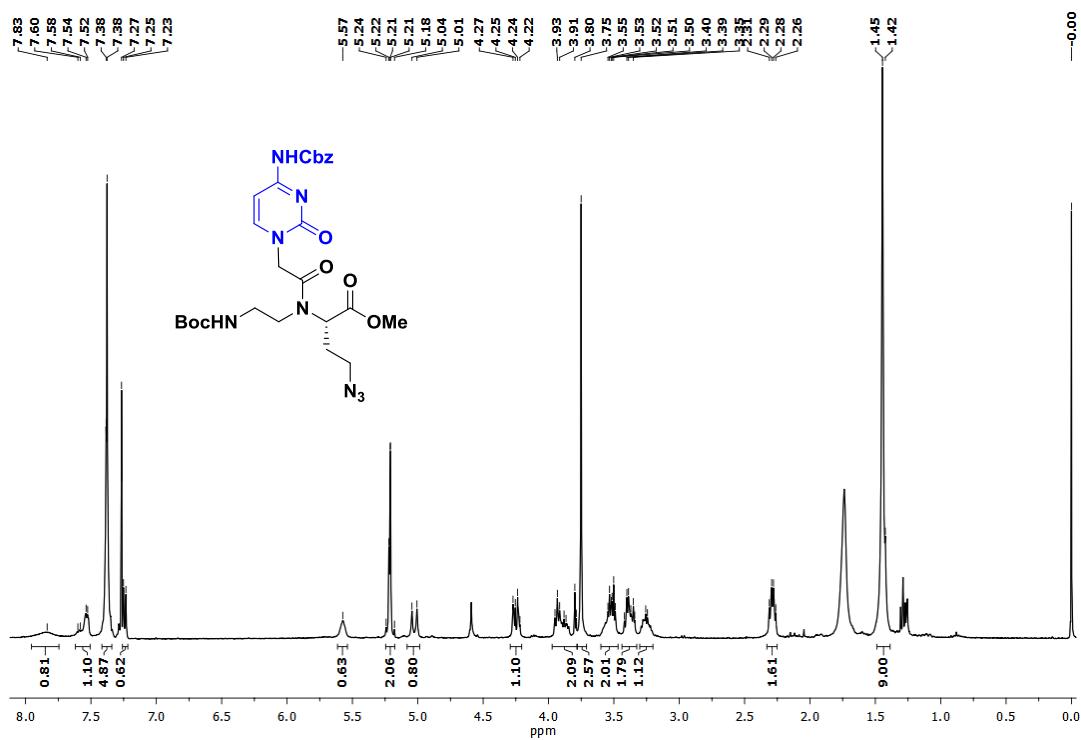
^{13}C NMR of Compound 22
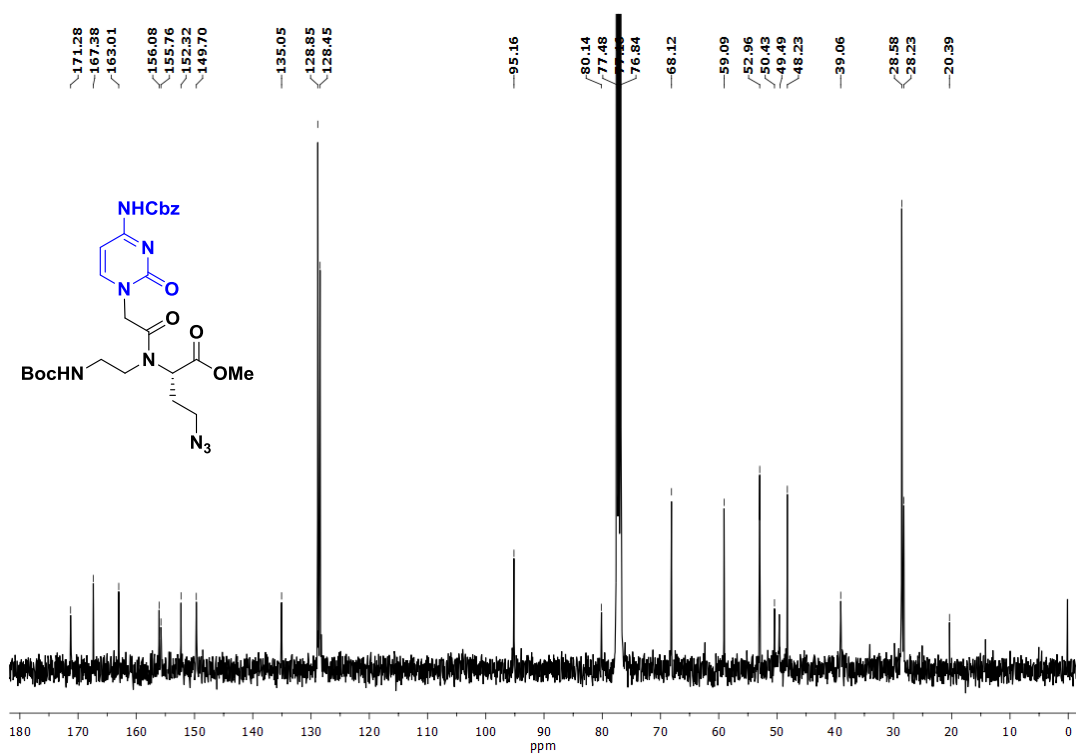
IR spectrum of Compound 22

HRMS Spectrum of Compound 22

Calculated mass $(M+H)^+ = 587.2578$

Obs. mass $(M+H)^+ = 587.2581$
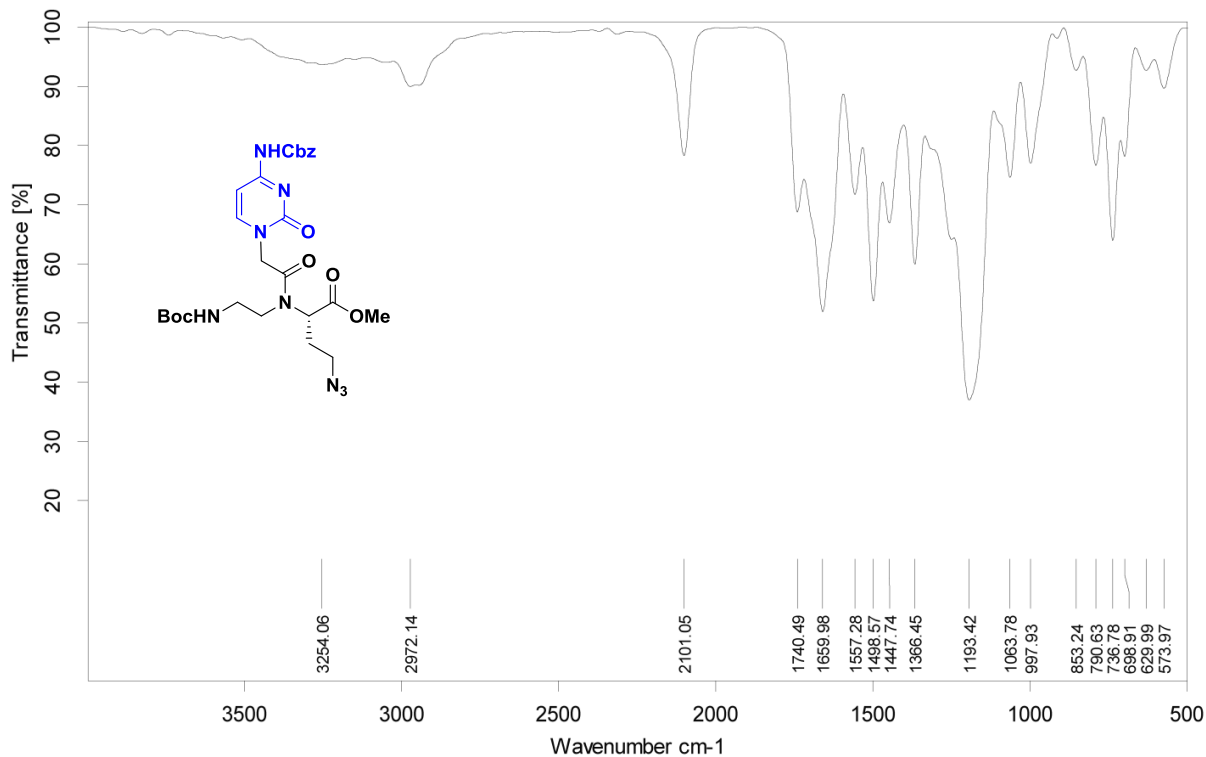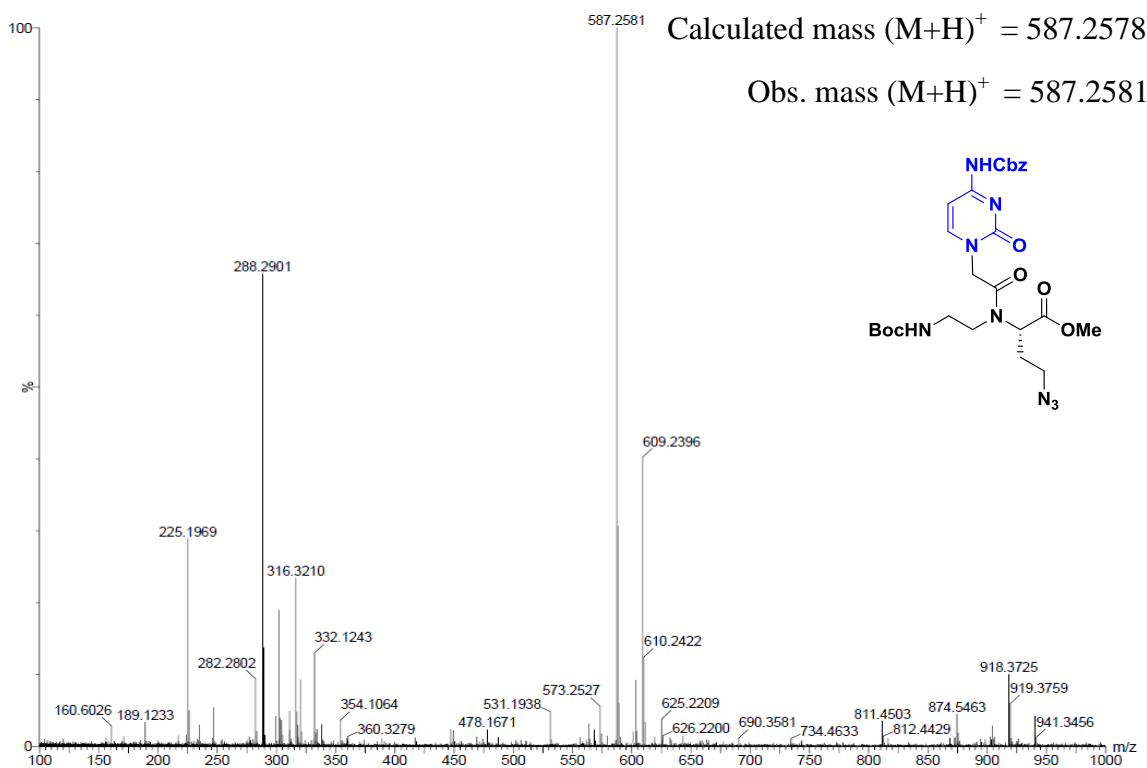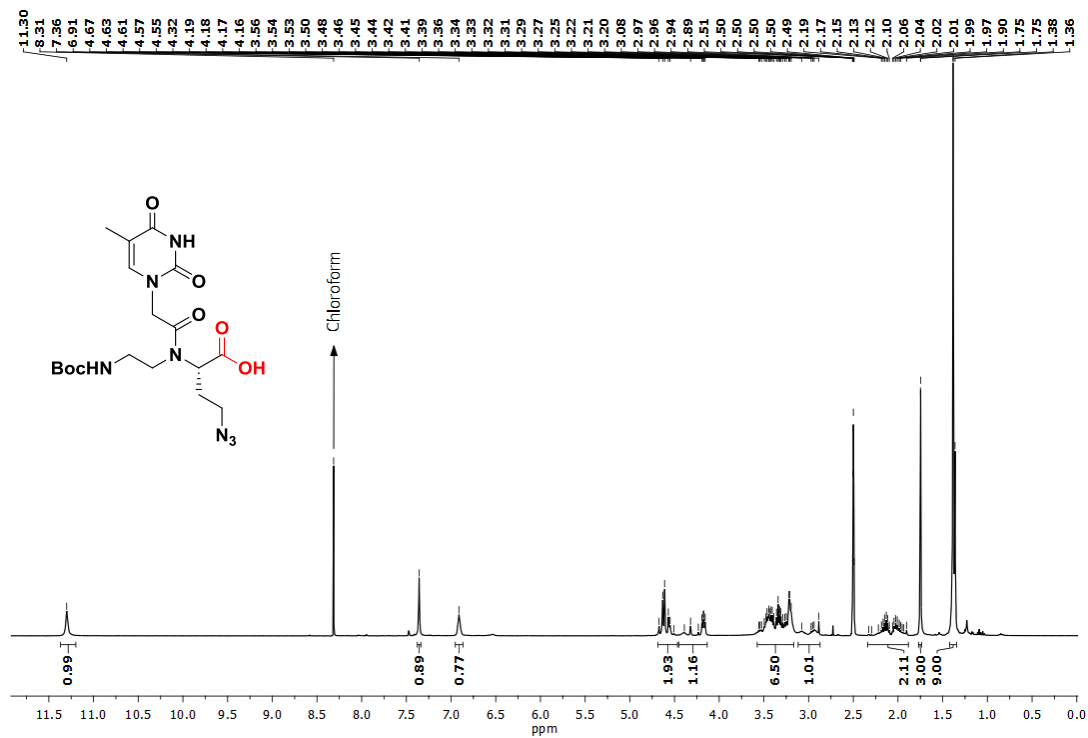
¹H NMR of Compound 4

¹³C NMR of Compound 4
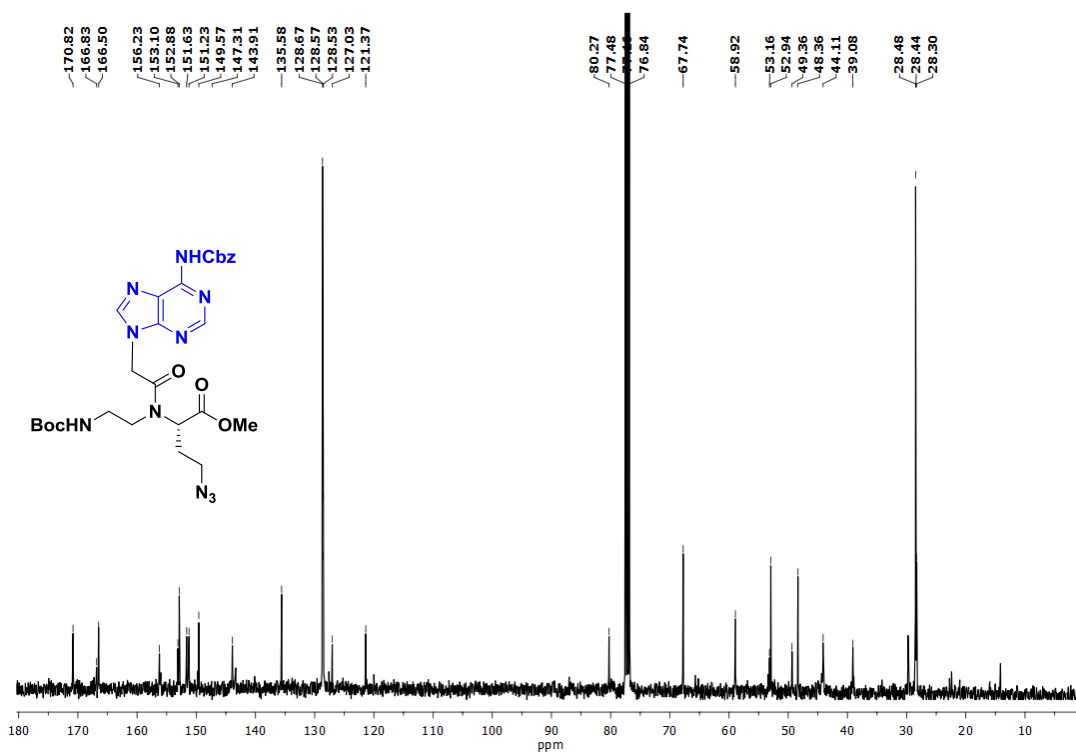
IR Spectrum of Compound 4

HRMS of Compound 4

Calculated mass (M+Na)$^+$ = 476.1870

Obs. mass (M+H)$^+$ = 476.1840
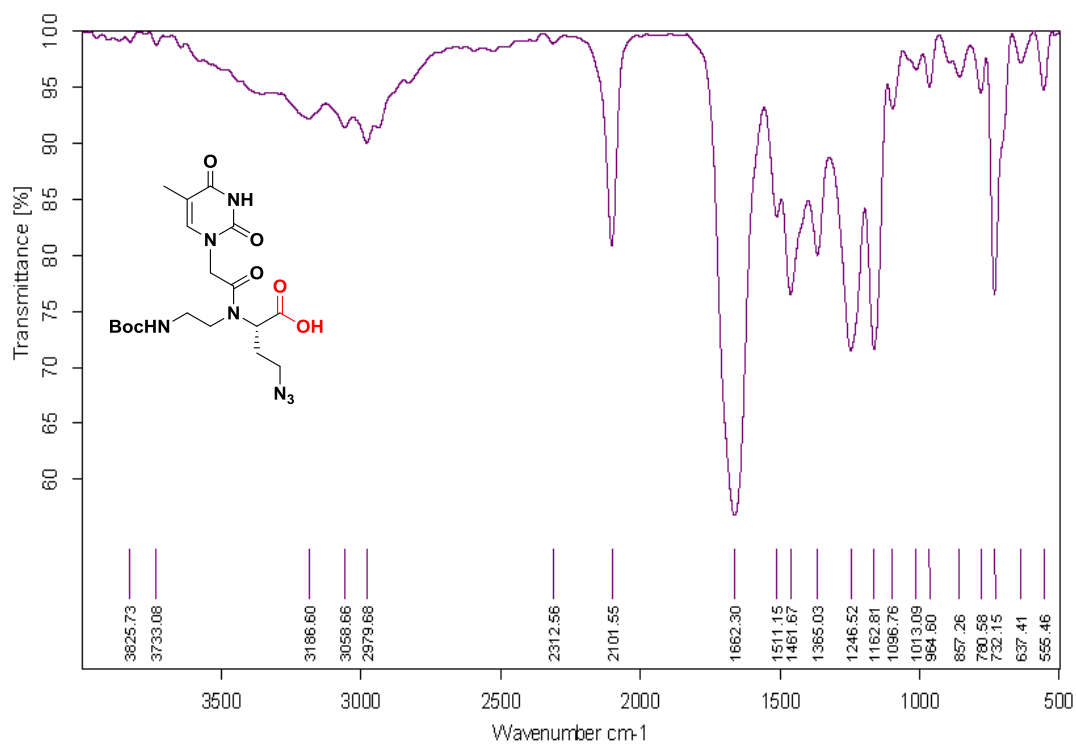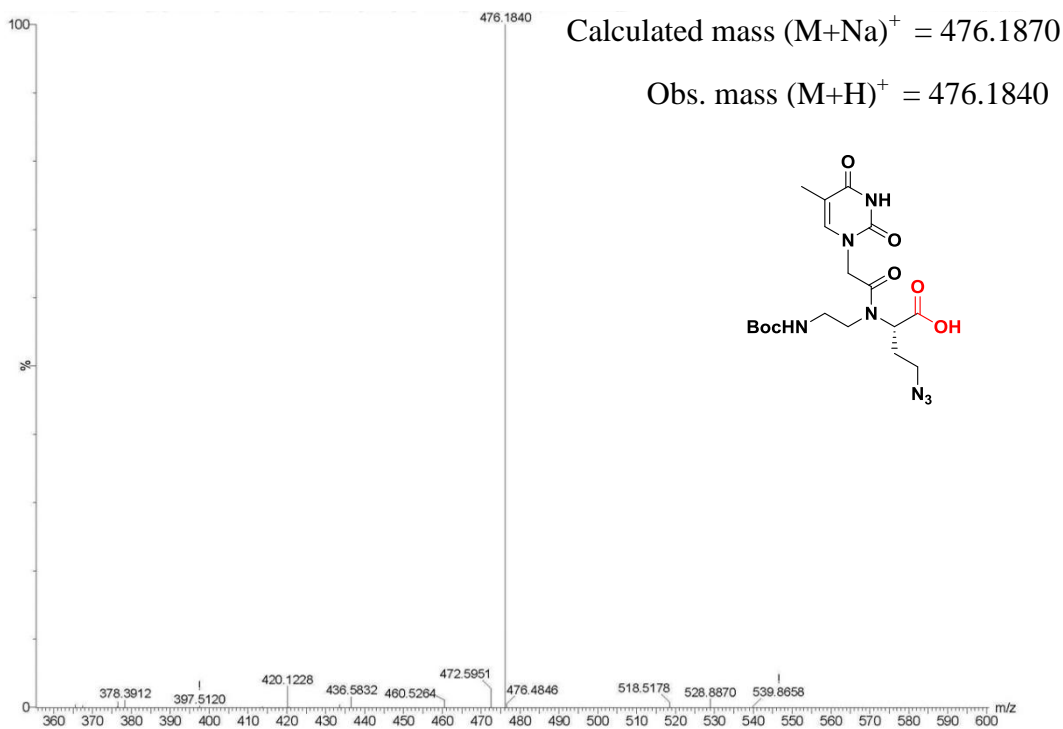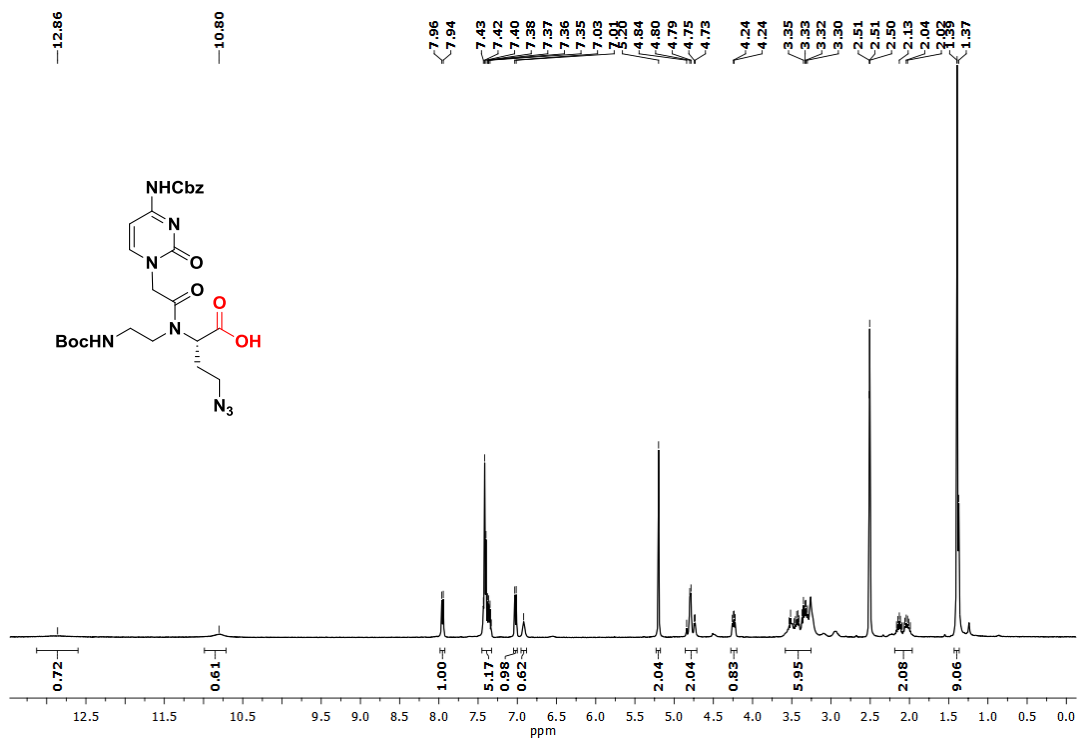
¹H NMR of Compound 5

¹³C NMR of Compound 5
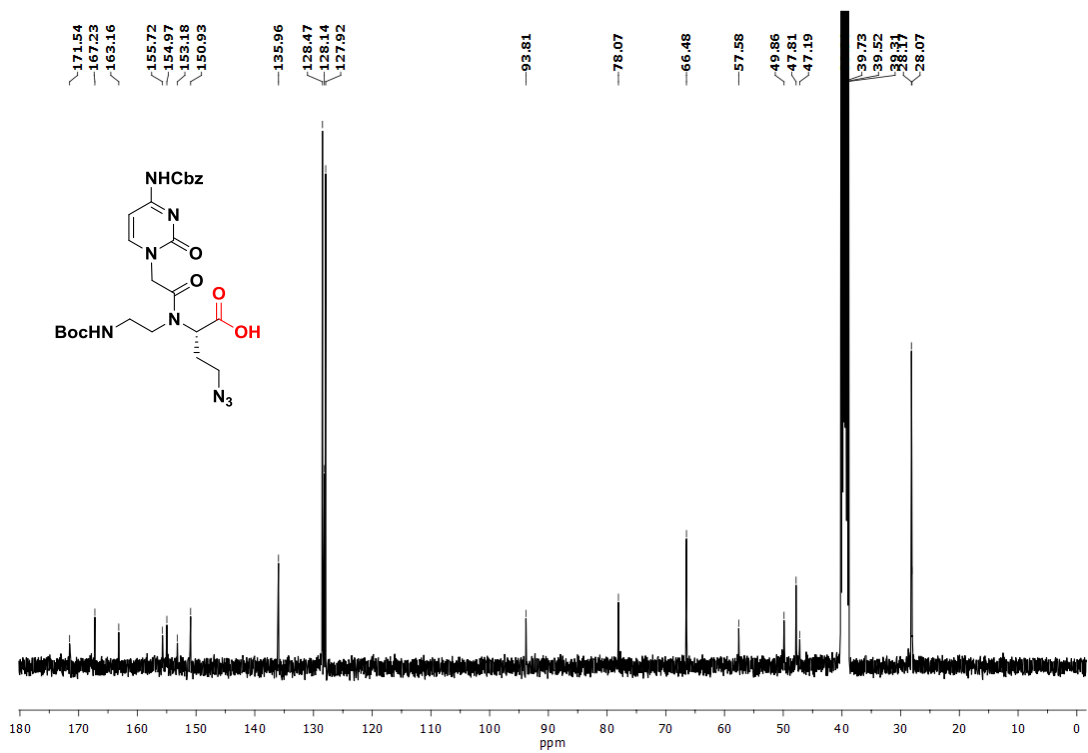
IR spectrum of Compound 5

HRMS of Compound 5

Calculated mass $(M+H)^+ = 573.2421$

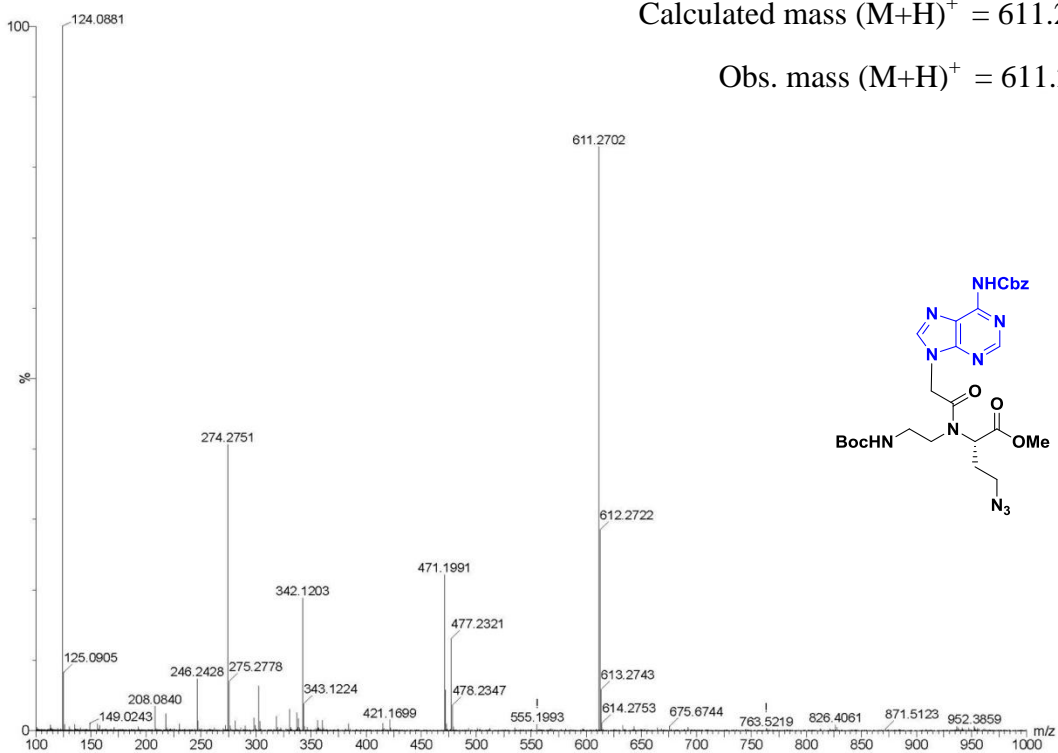Obs. mass $(M+H)^+ = 573.2424$
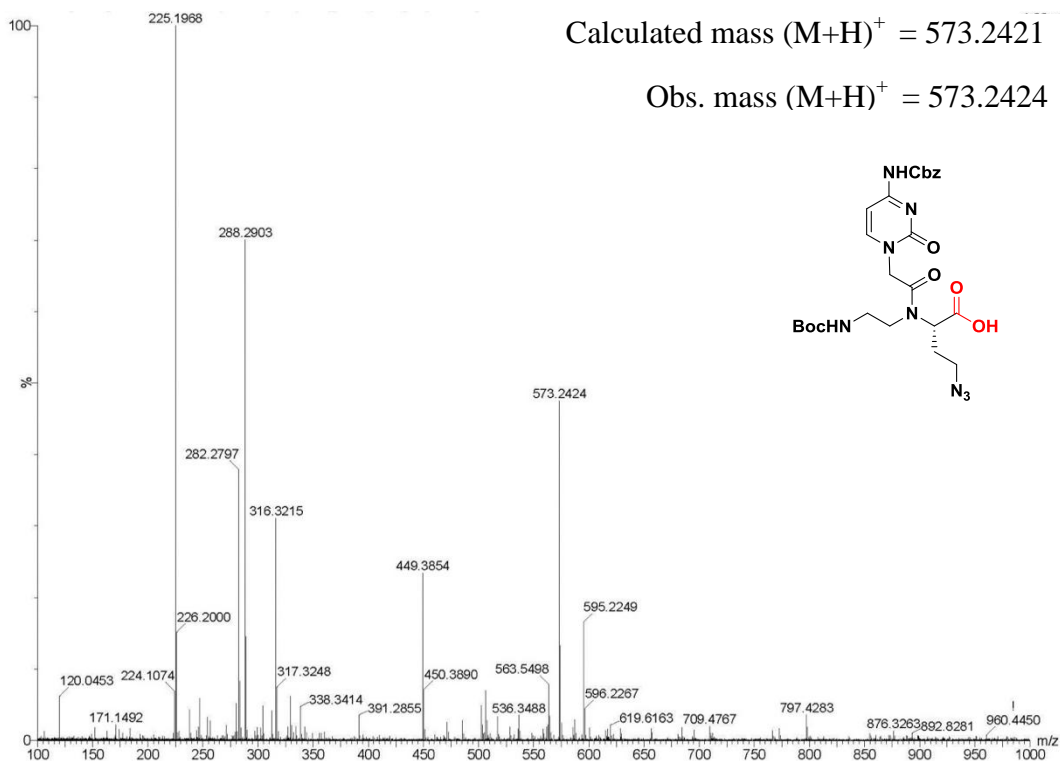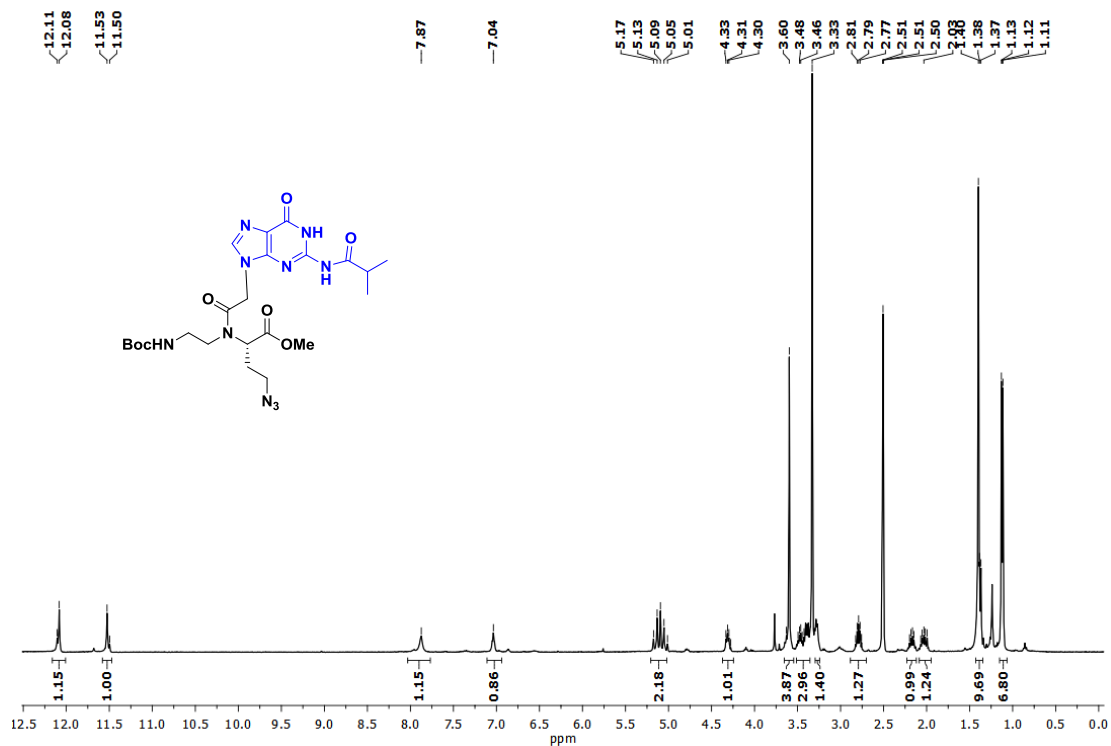
¹H NMR of Compound 23
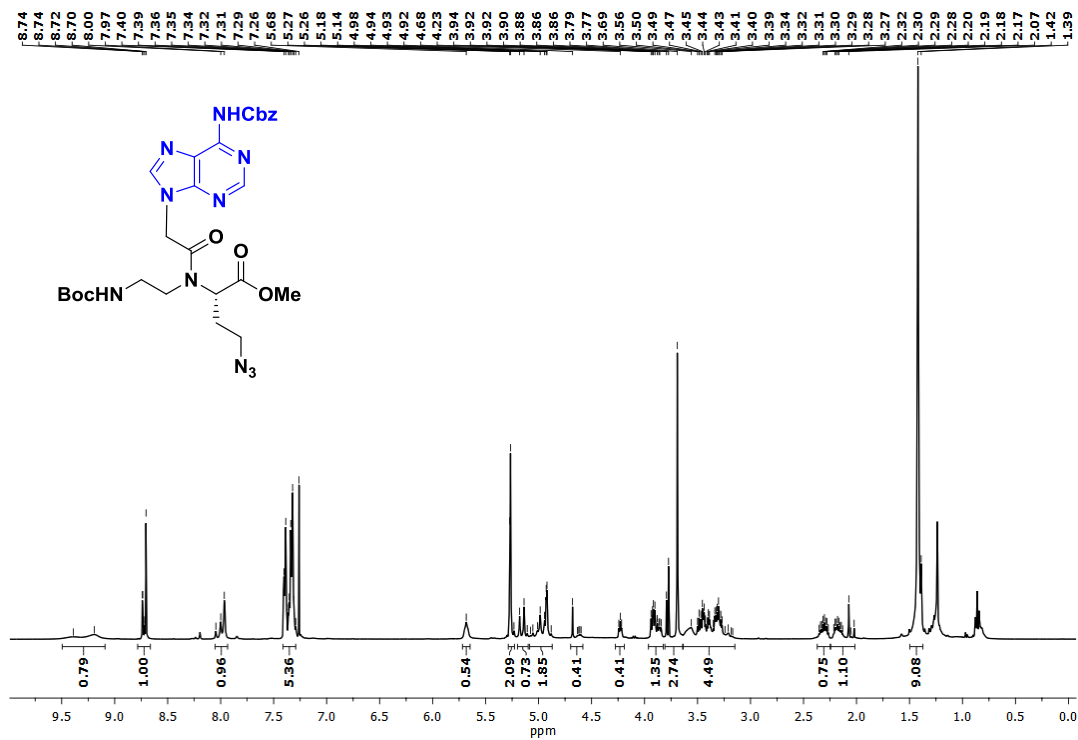
¹³C NMR of Compound 23

IR spectrum of Compound 23

HRMS of Compound 23

Calculated mass $(M+H)^+ = 611.2690$

Obs. mass $(M+H)^+ = 611.2702$
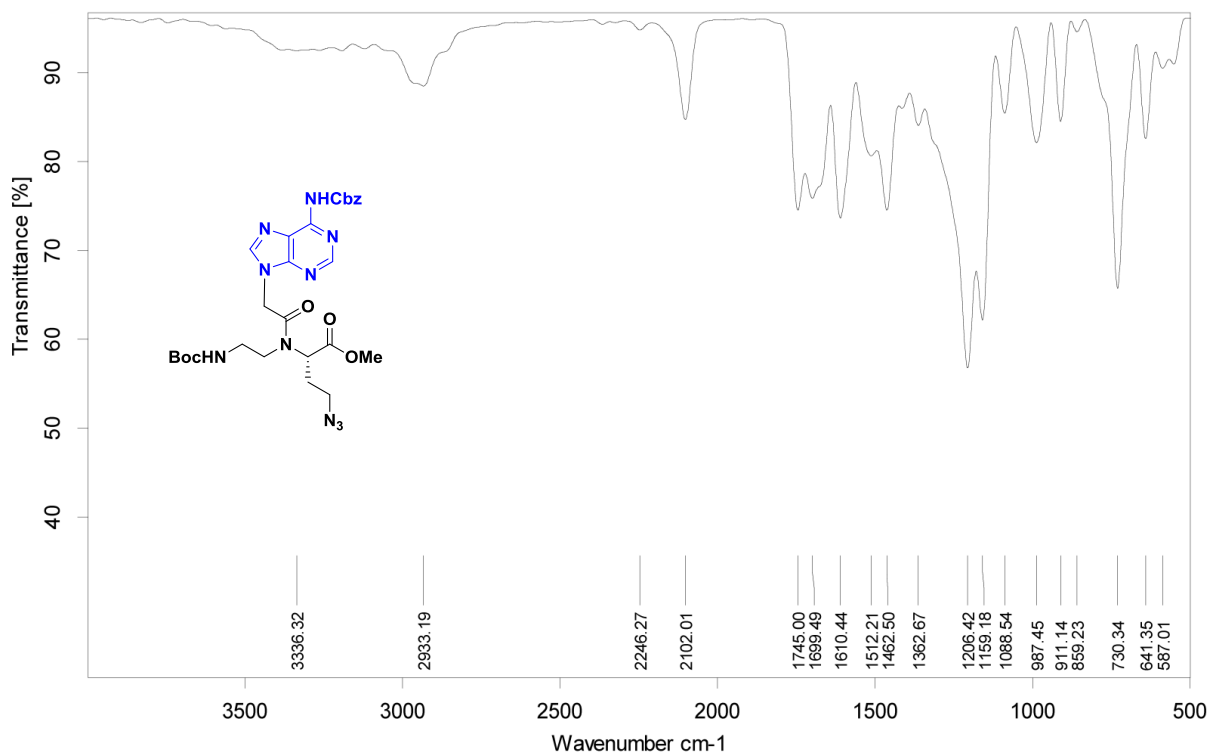
¹H NMR of Compound 24

¹³C NMR of Compound 24

IR Spectrum of Compound 24

HRMS of Compound 24

Calculated mass $(M+H)^+ = 563.2690$

Obs. mass $(M+H)^+ = 563.2678$
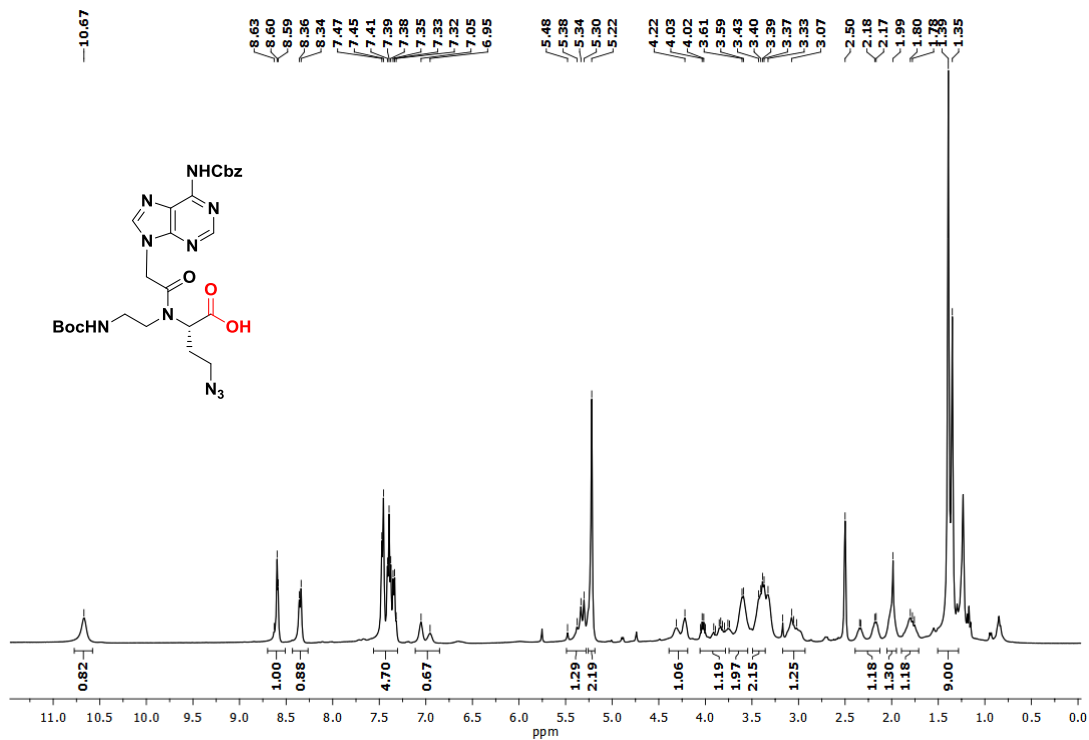
¹H NMR of Compound 6

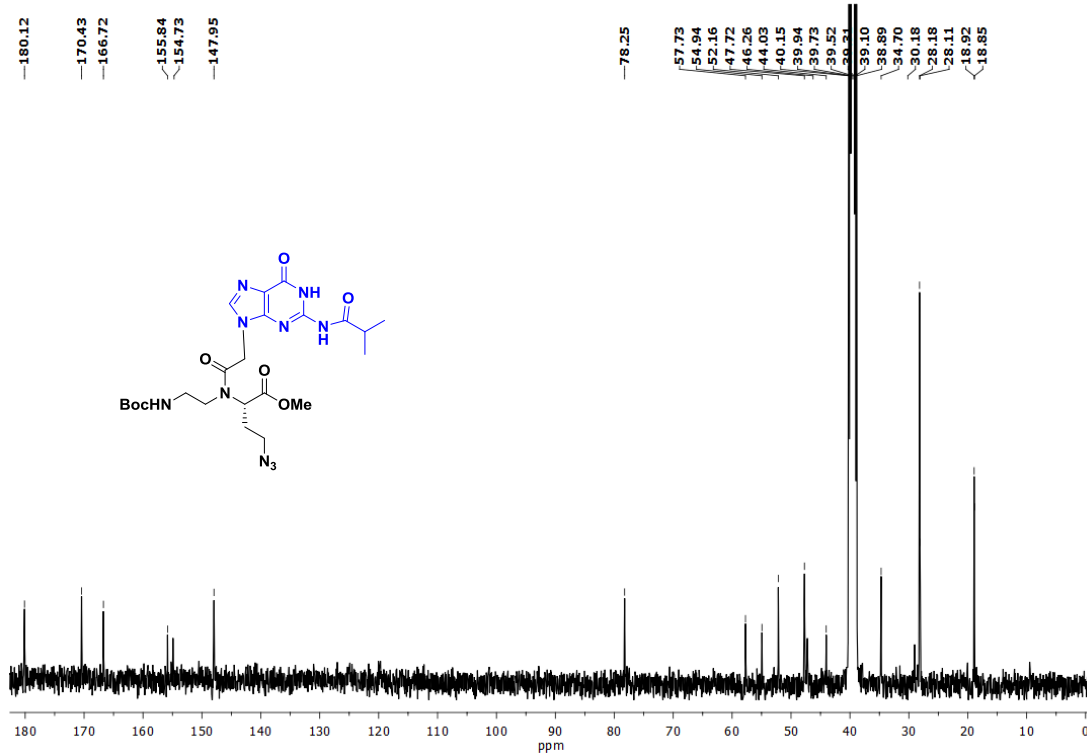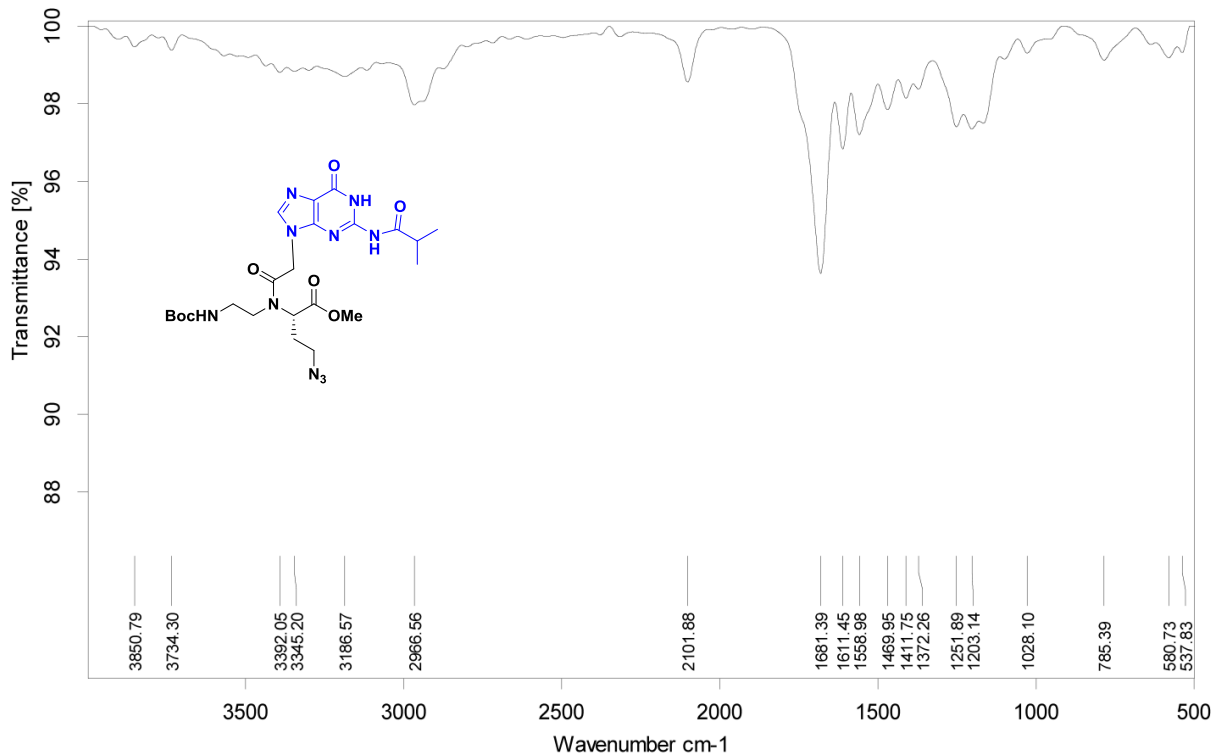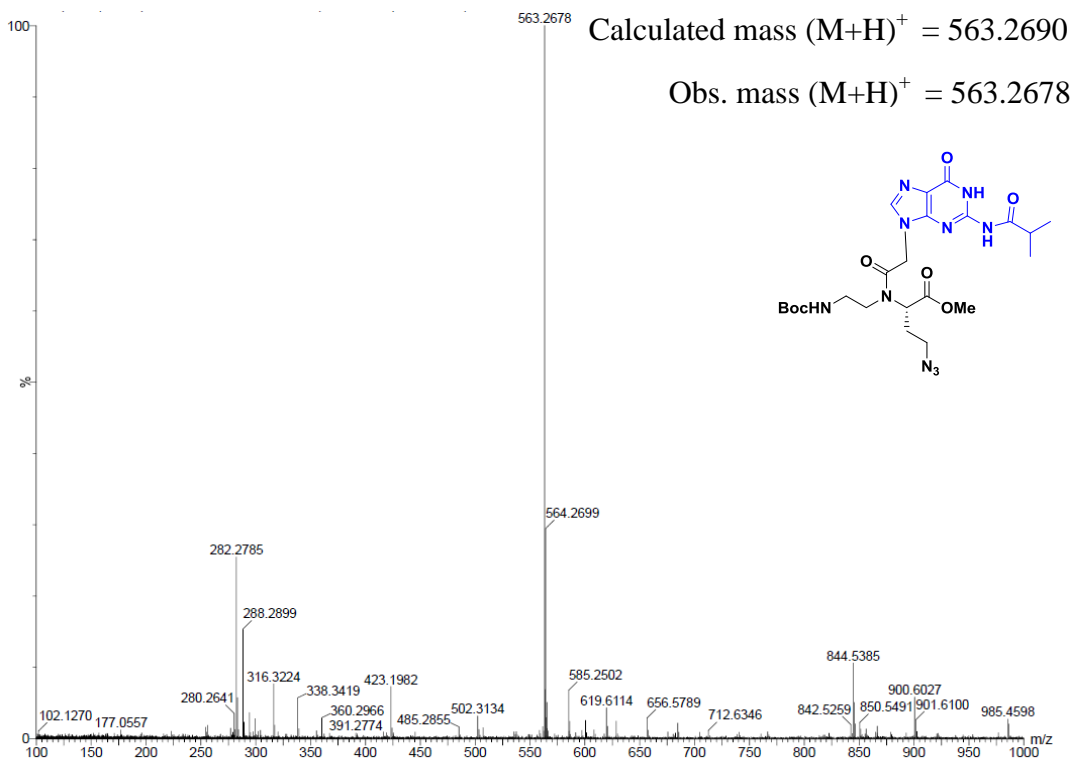

¹³C NMR of Compound 6

IR Spectrum of Compound 6

HRMS of Compound 6

Calculated mass $(M+H)^+ = 597.2534$

Obs. mass $(M+H)^+ = 597.2524$
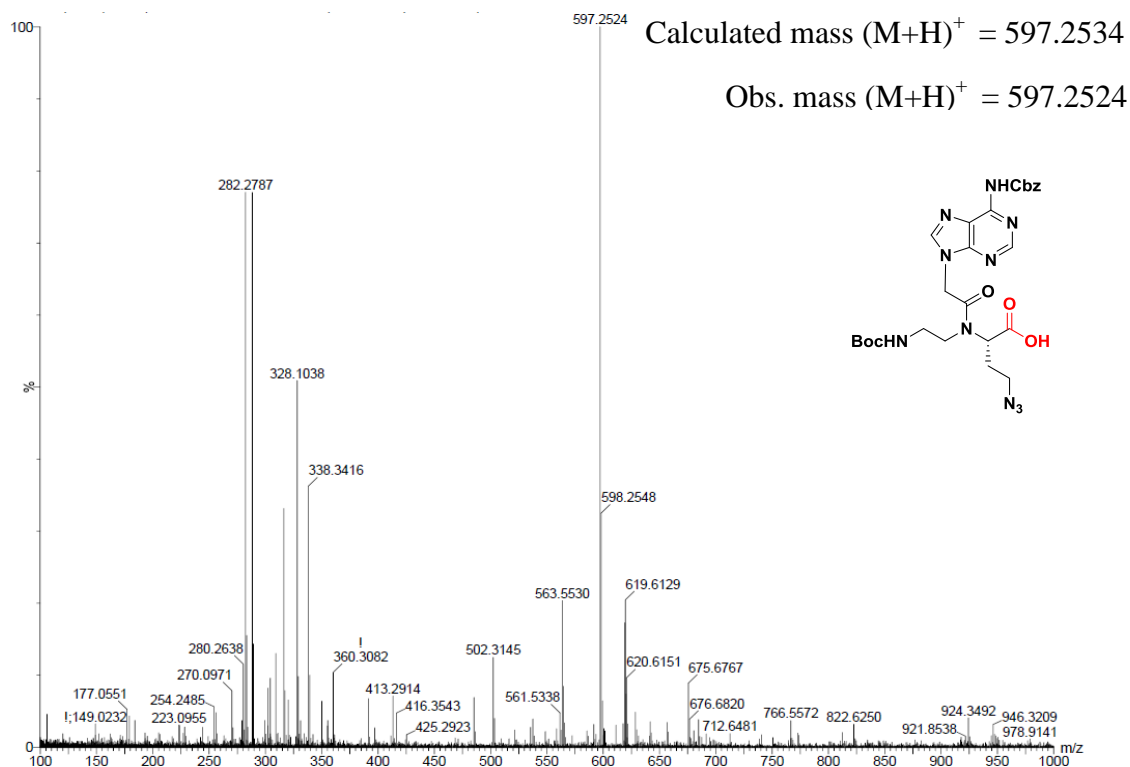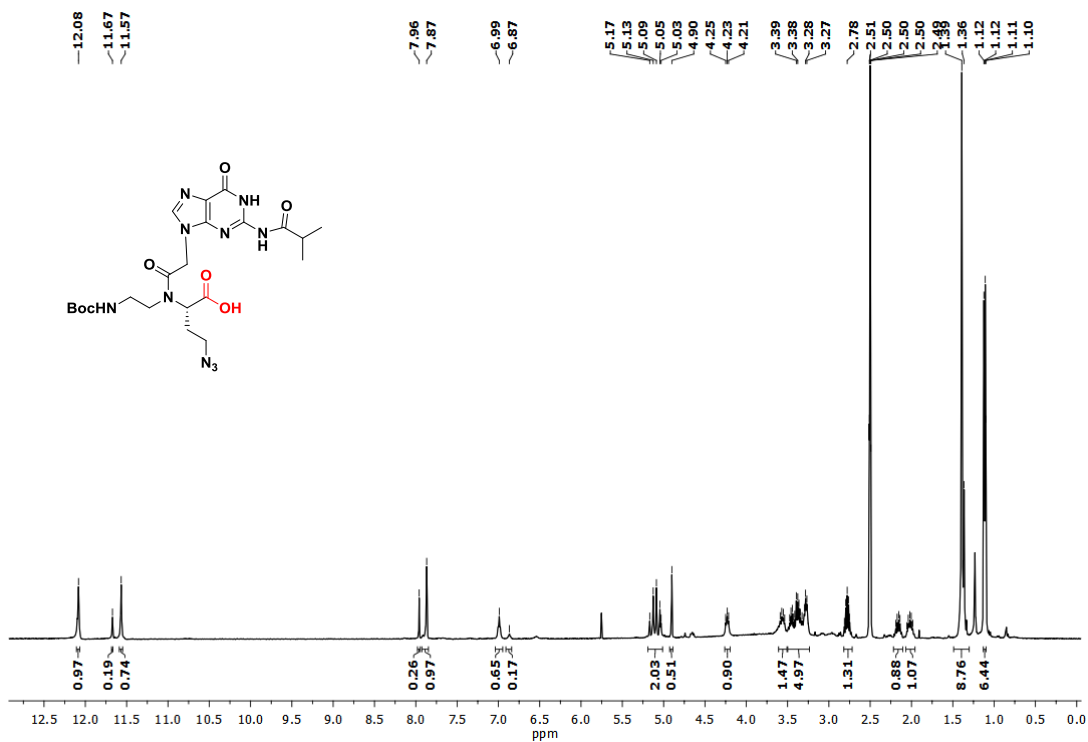
^{1}H NMR of Compound 7

^{13}C NMR of Compound 7
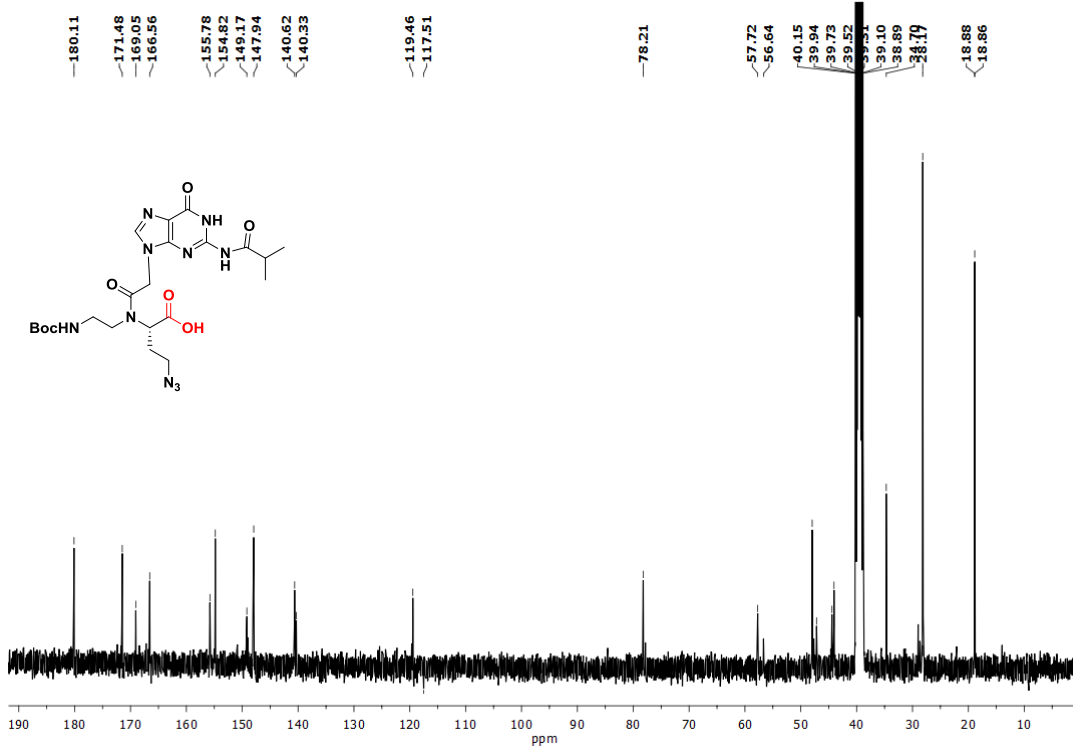
IR Spectrum of Compound 7

HRMS of Compound 7

Calculated mass $(M+H)^+ = 549.2534$

Obs. mass $(M+H)^+ = 549.2535$
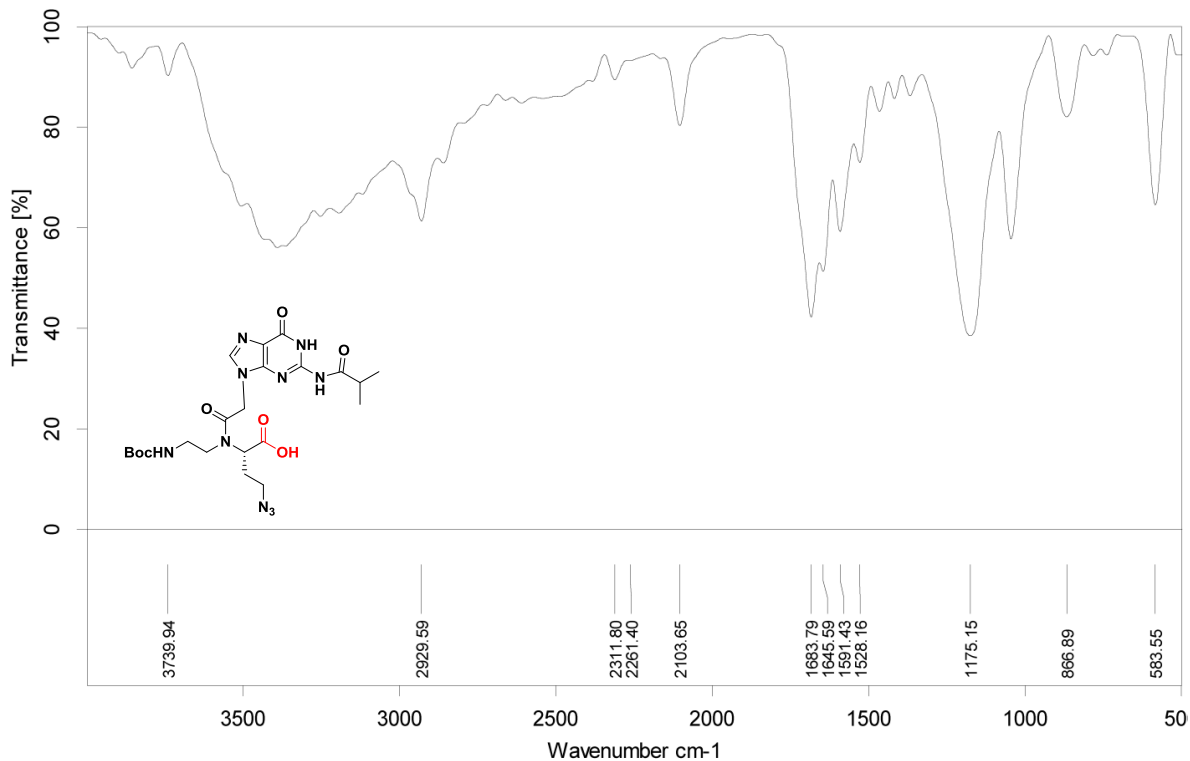
^{1}H NMR of Compound 25
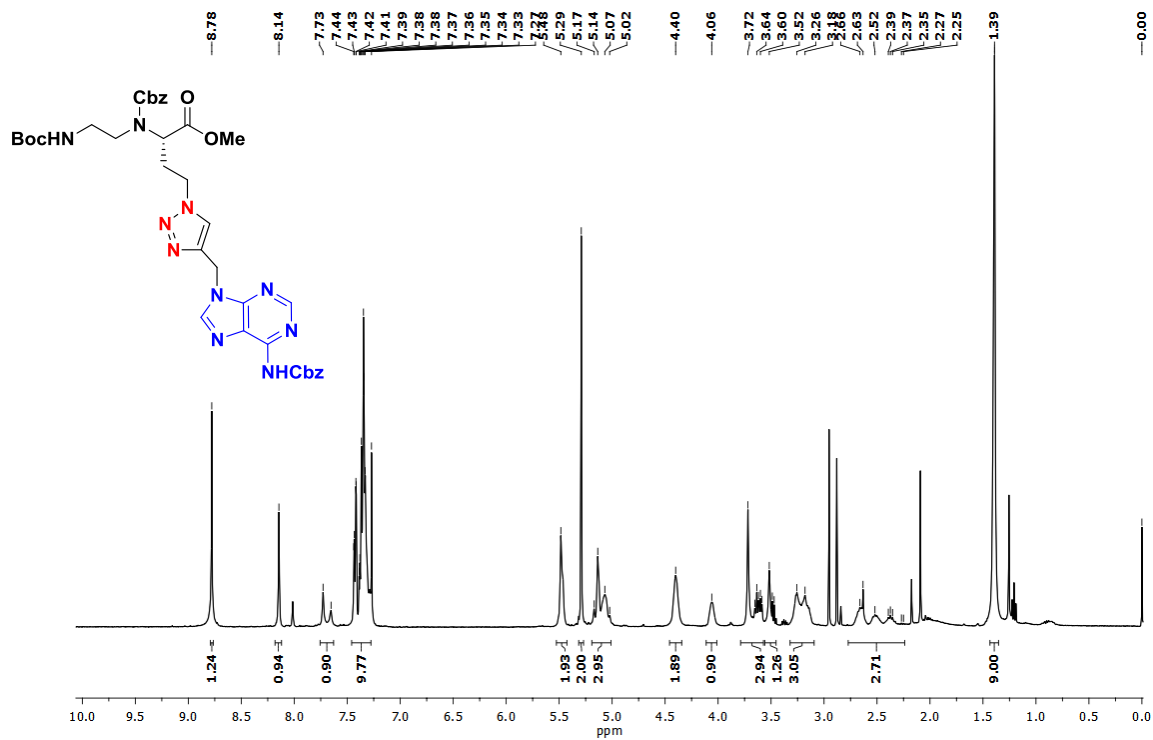
^{13}C NMR of Compound 25
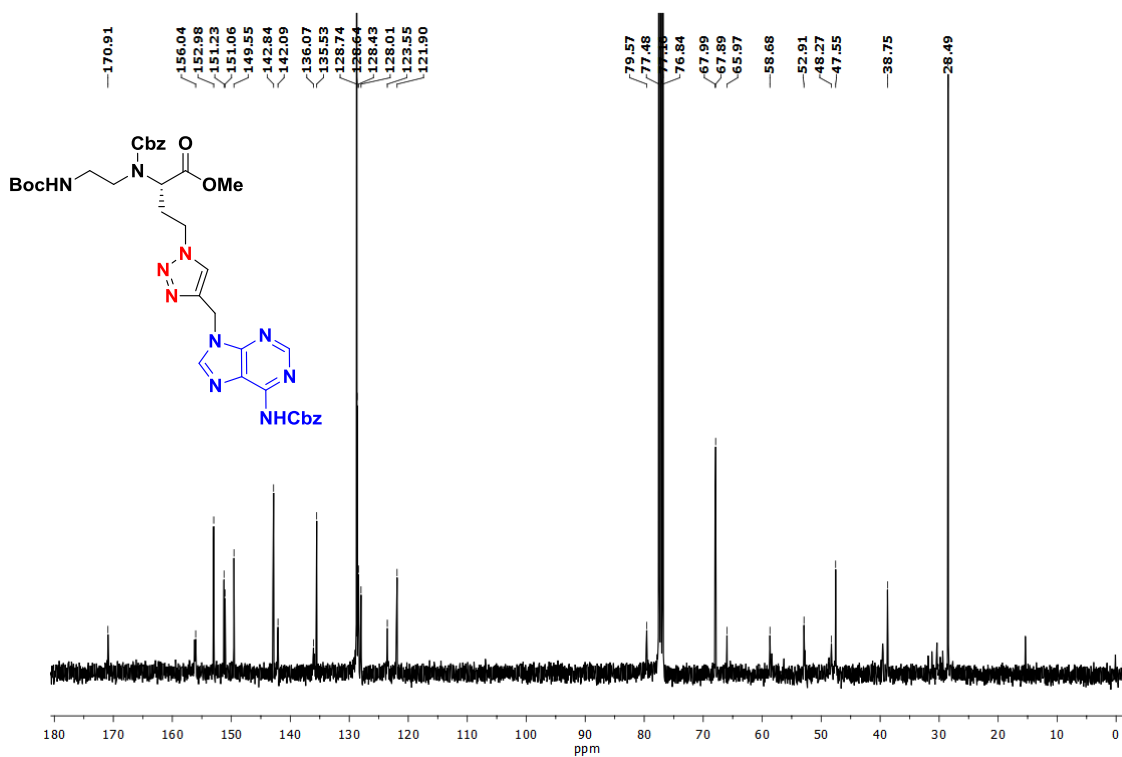
^{1}H NMR of Compound 9
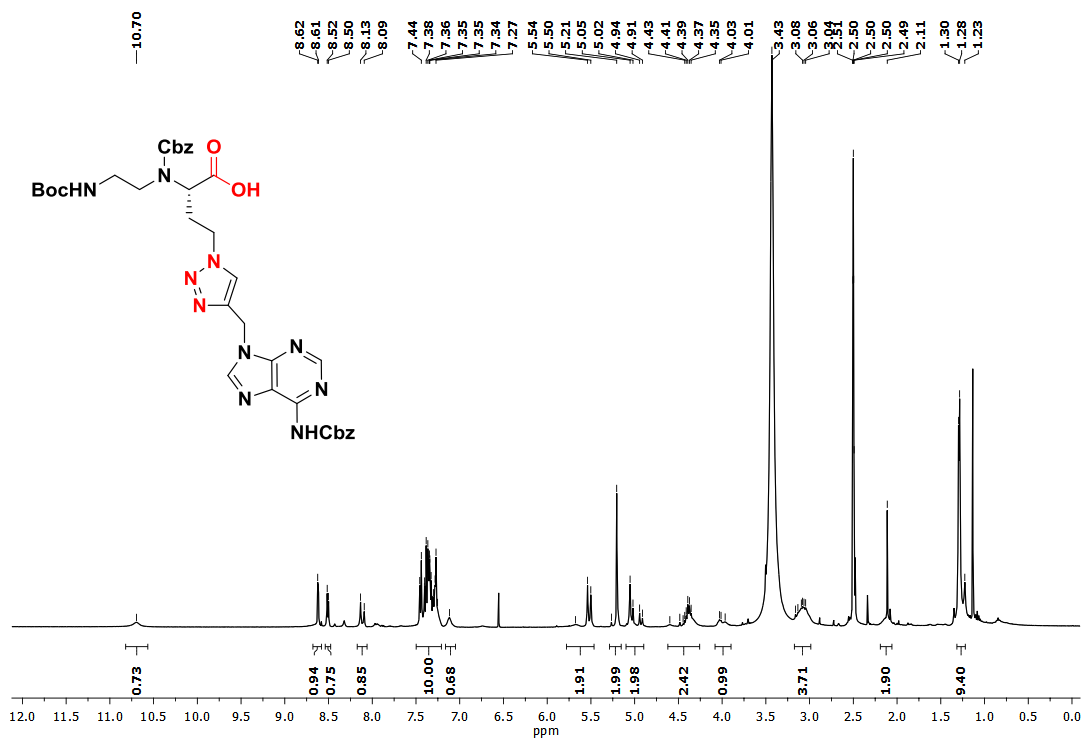
^{13}C NMR of Compound 9
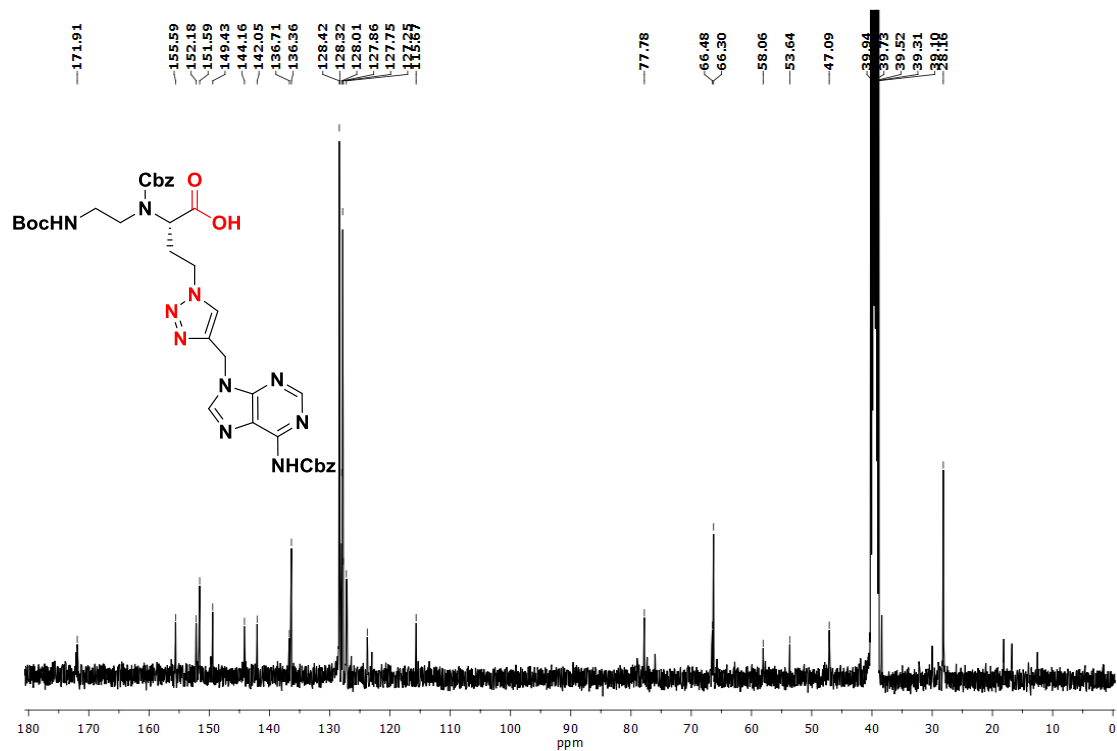
^{1}H NMR of Compound 26
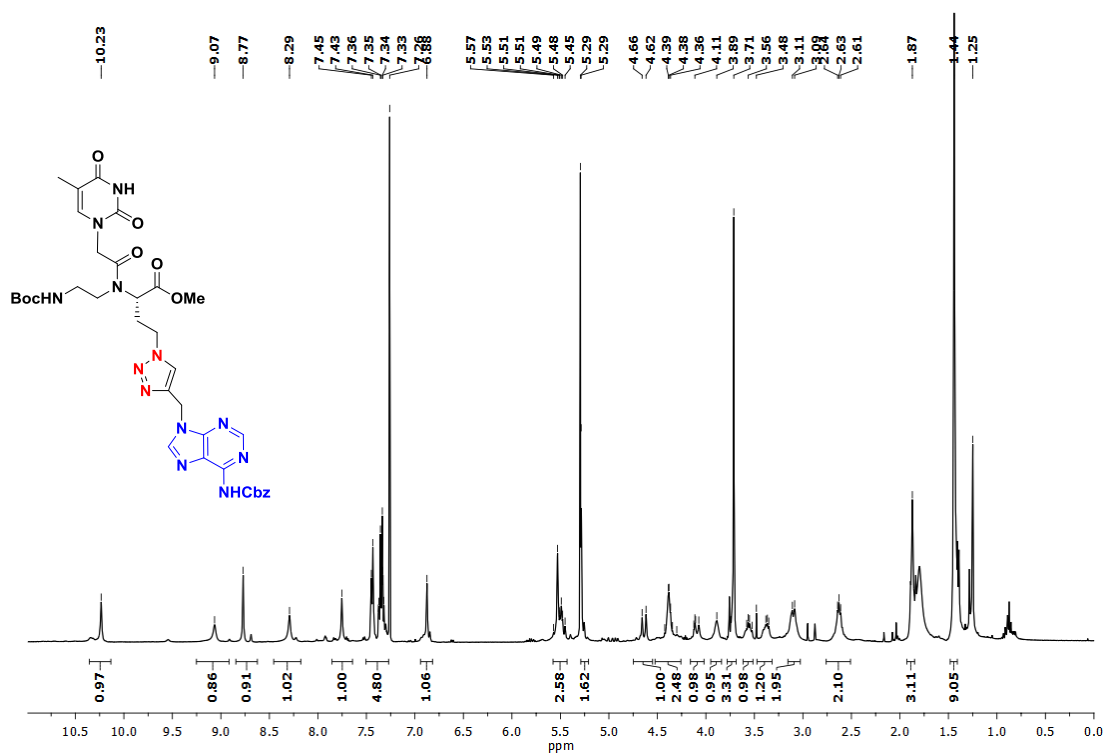
^{13}C NMR of Compound 26
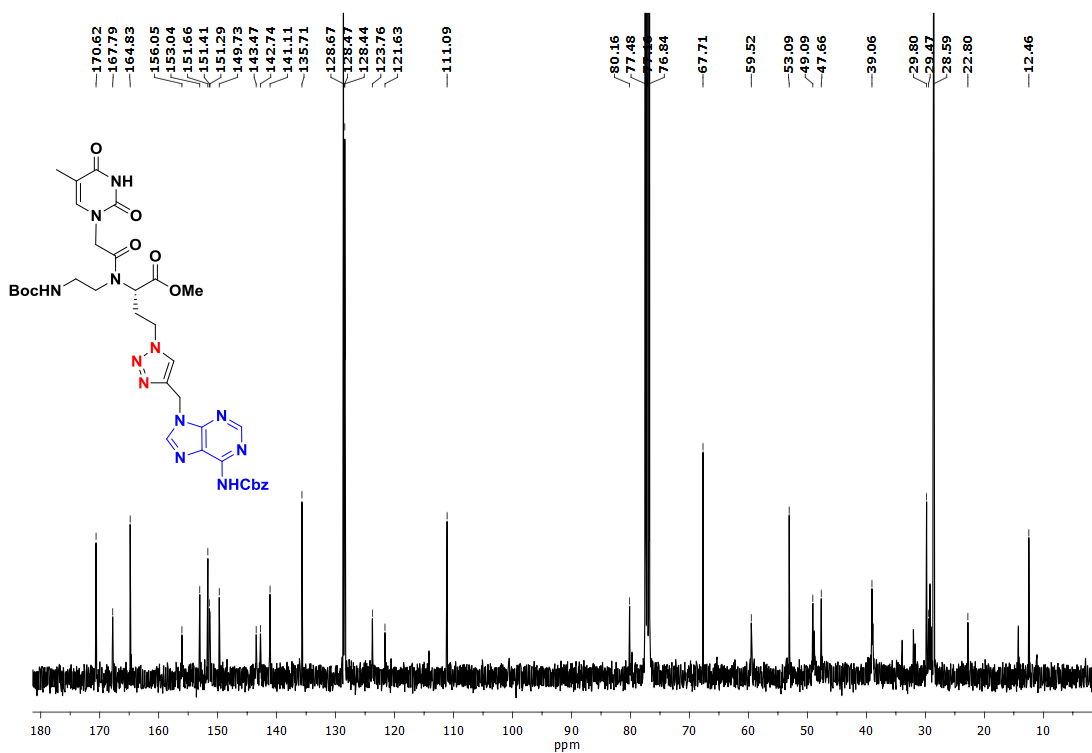
IR Spectrum of Compound 26

HRMS of Compound 26

Calculated mass $(M+H)^+ = 775.3276$

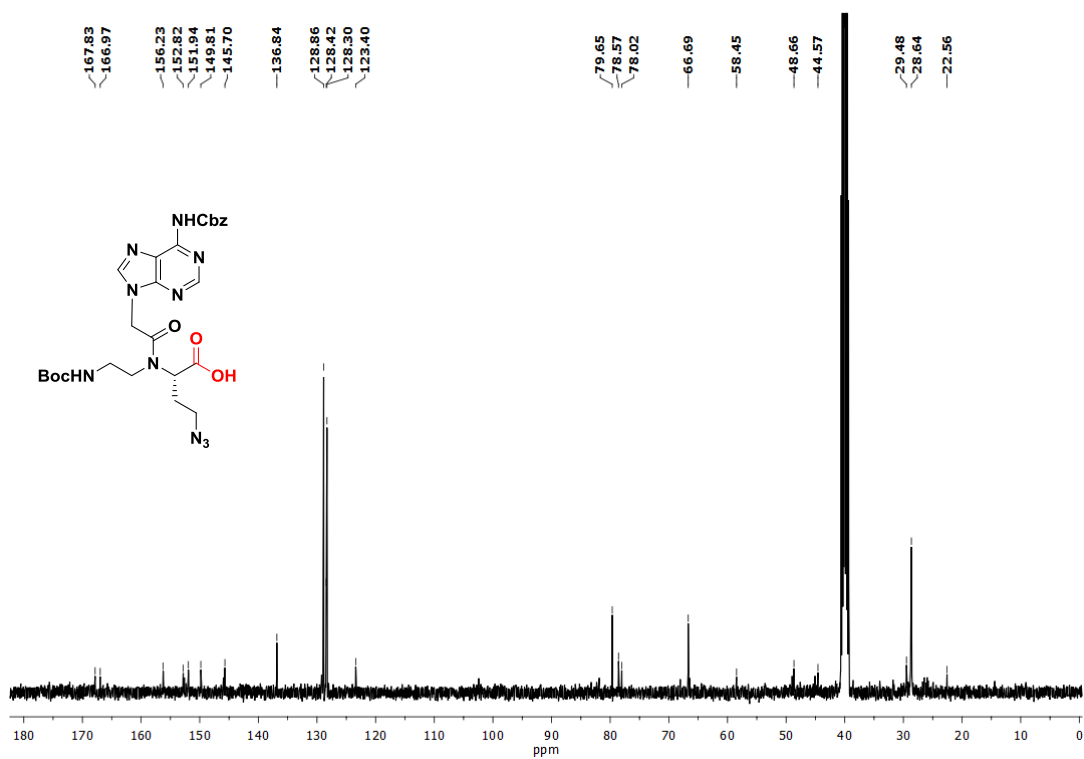Obs. mass $(M+H)^+ = 775.3275$
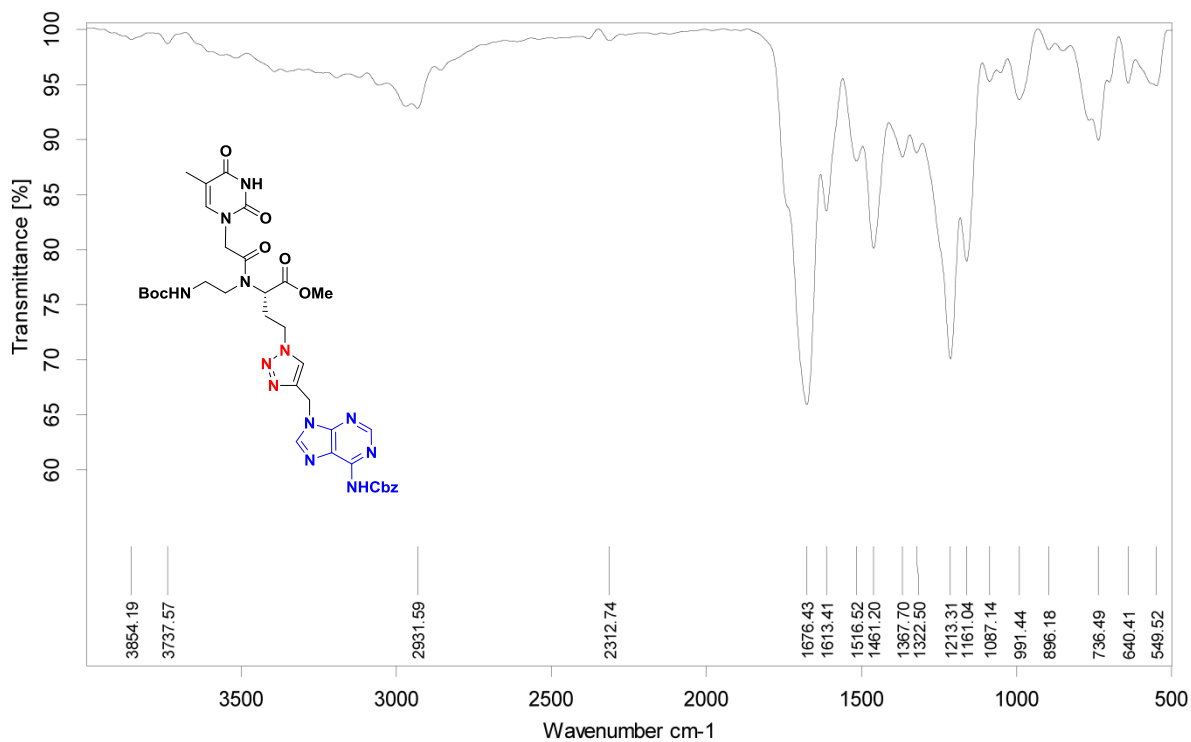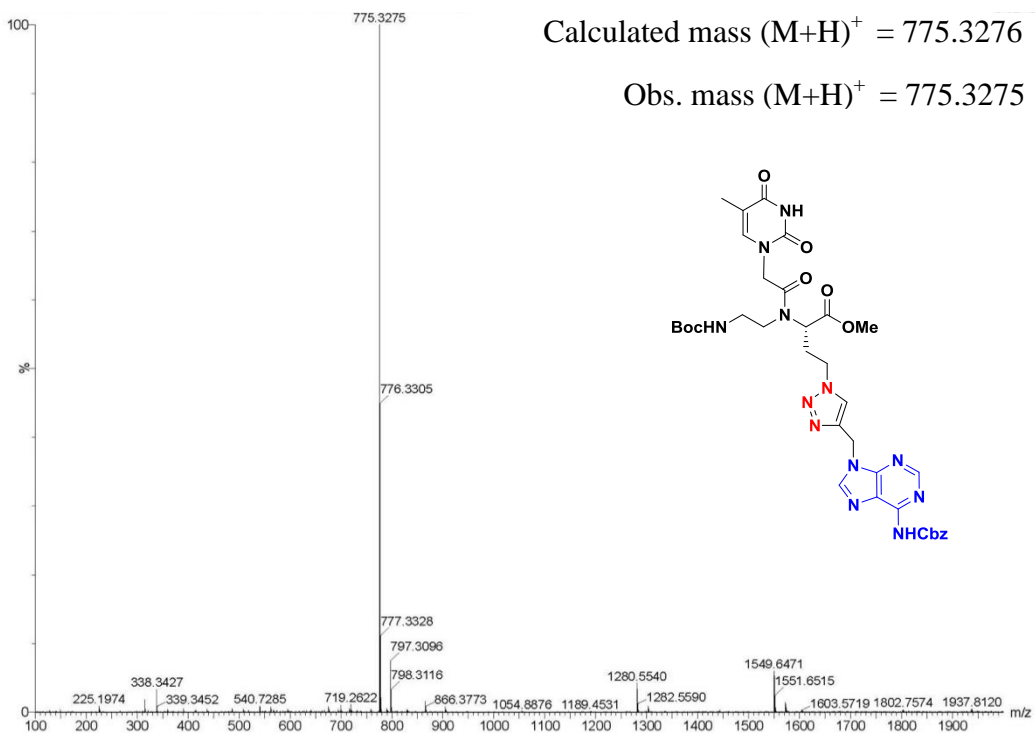
¹H NMR of Compound 10
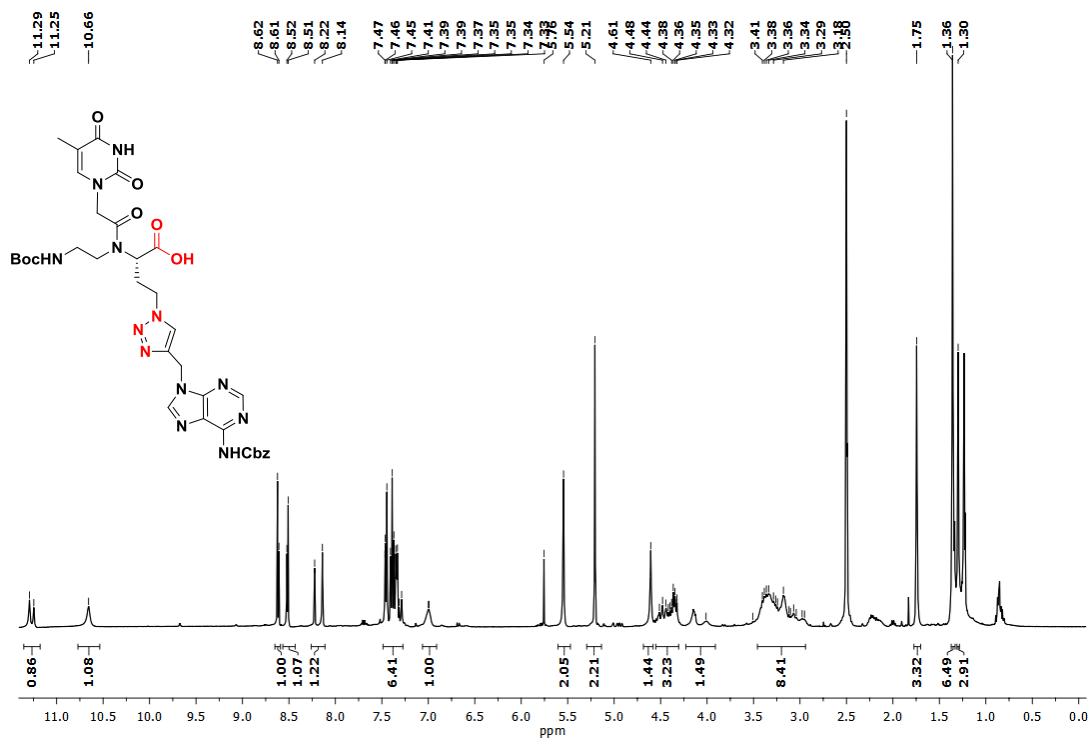
¹³C NMR of Compound 10
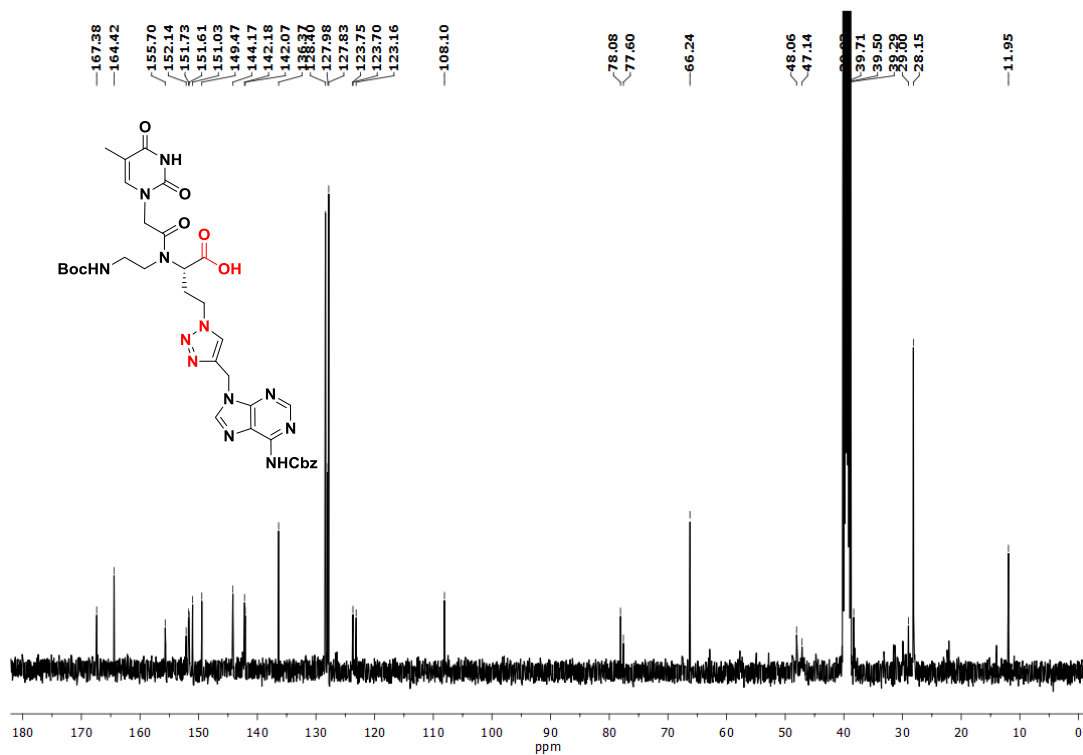
IR Spectrum of Compound 10

HRMS of Compound 10

Calculated mass $(M+H)^+ = 761.3119$

Obs. mass $(M+H)^+ = 761.3113$
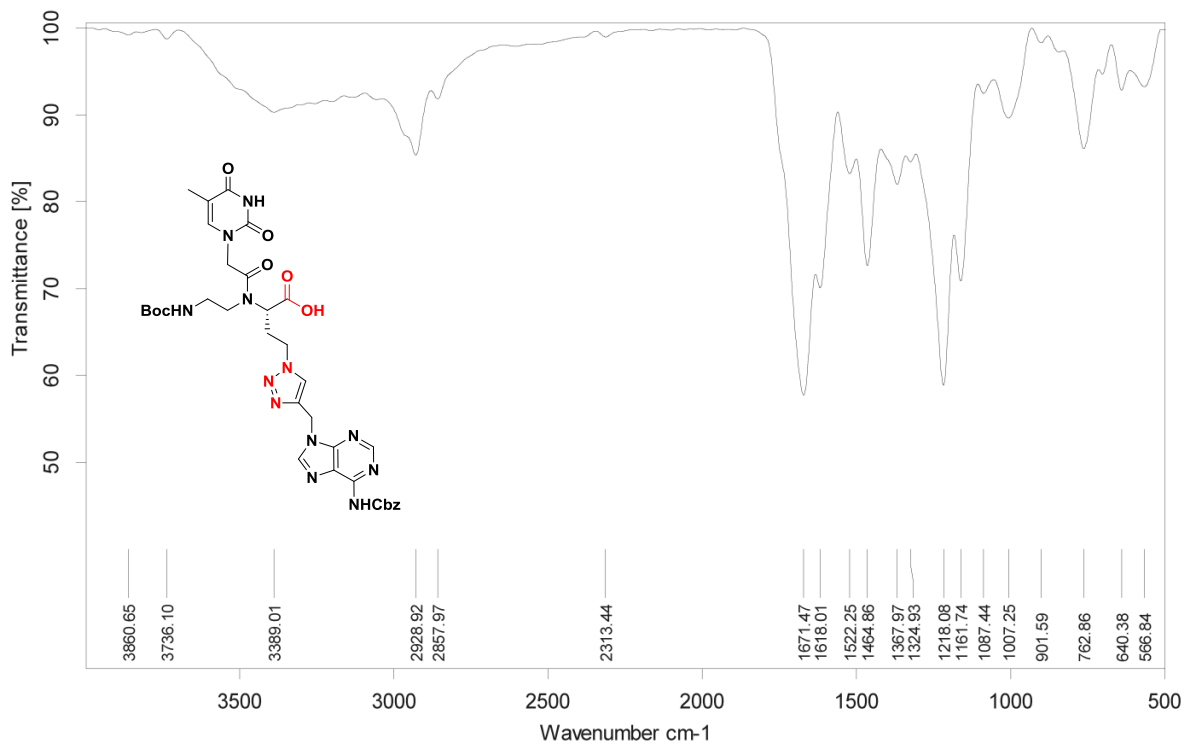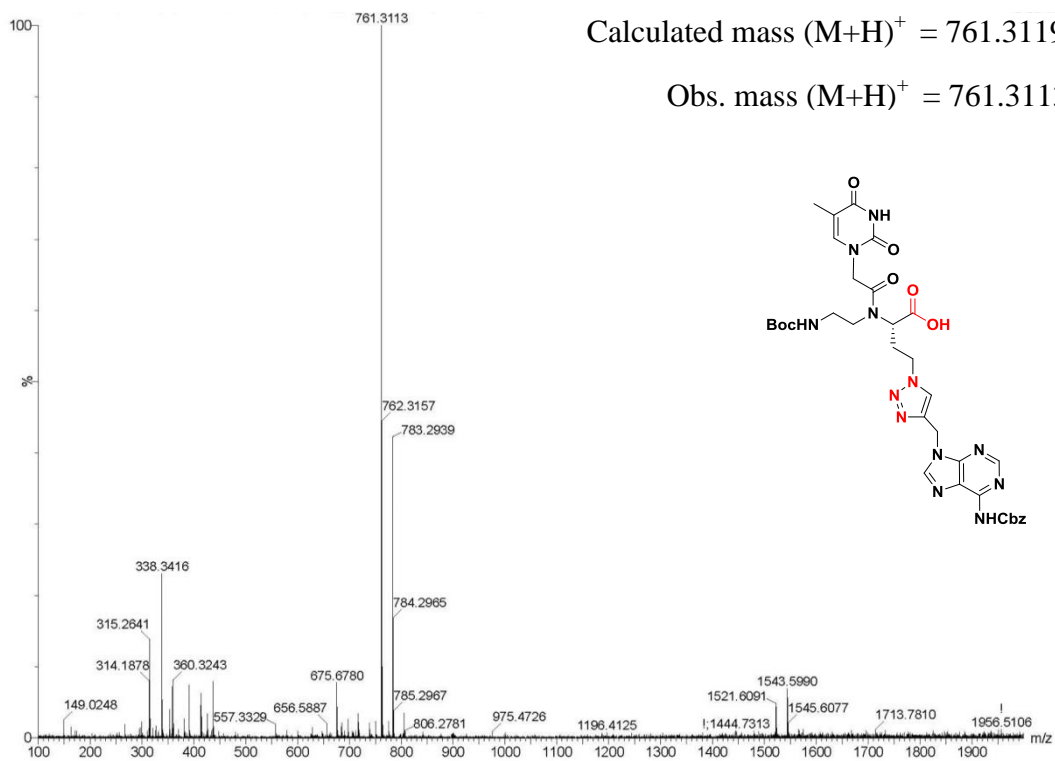
¹H NMR of Compound 29
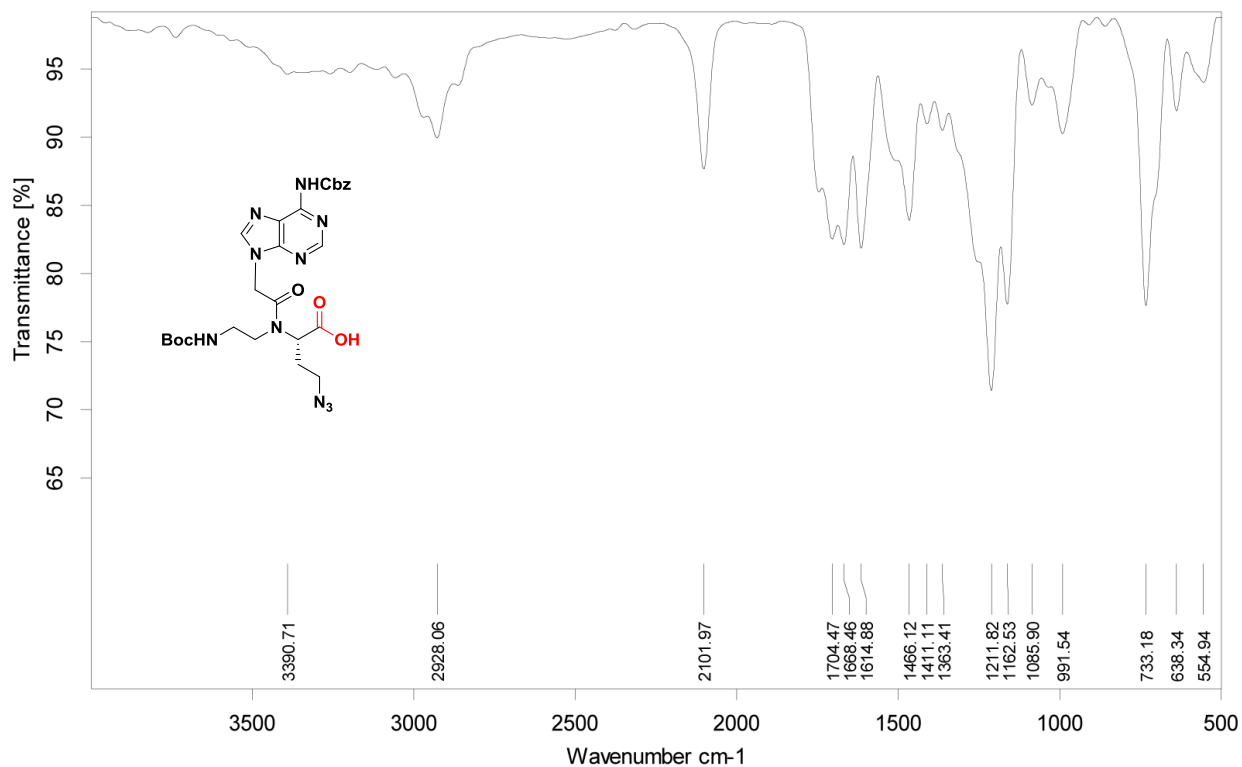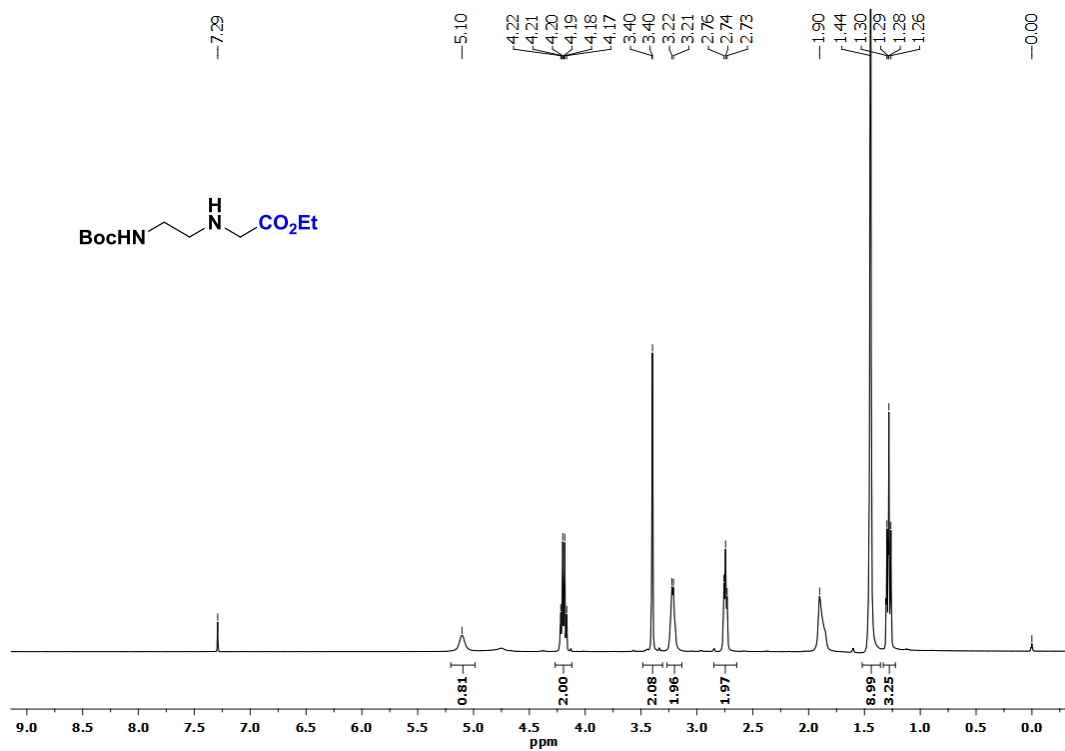

¹³C NMR of Compound 29
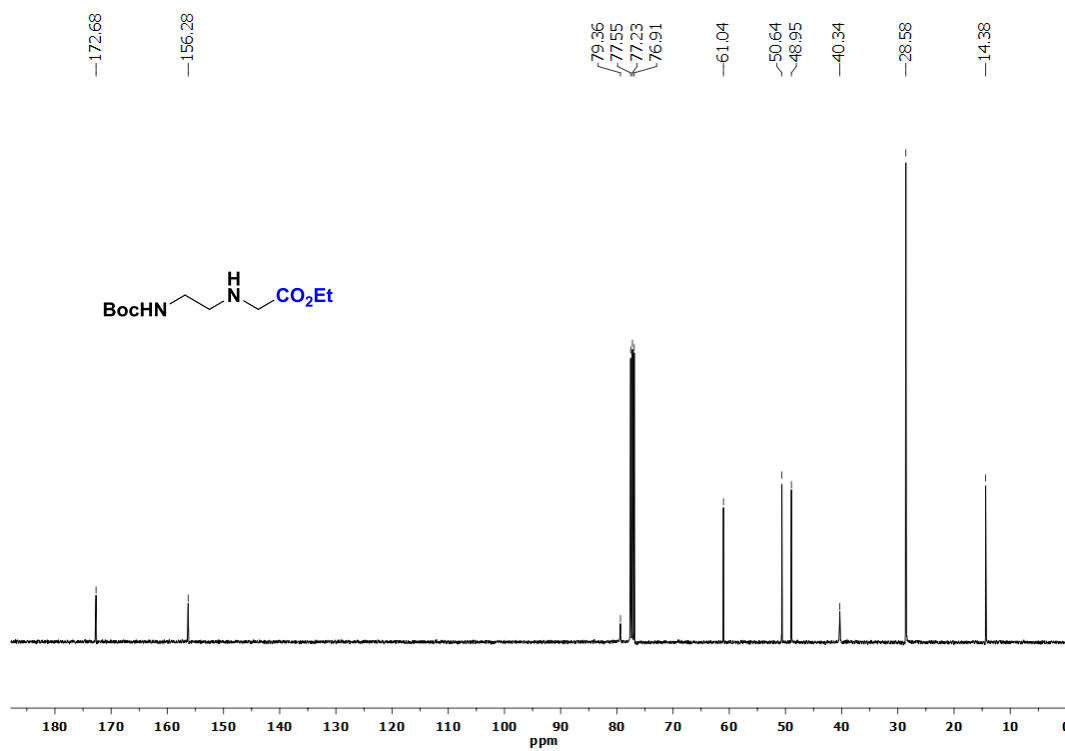
IR Spectrum of Compound 29
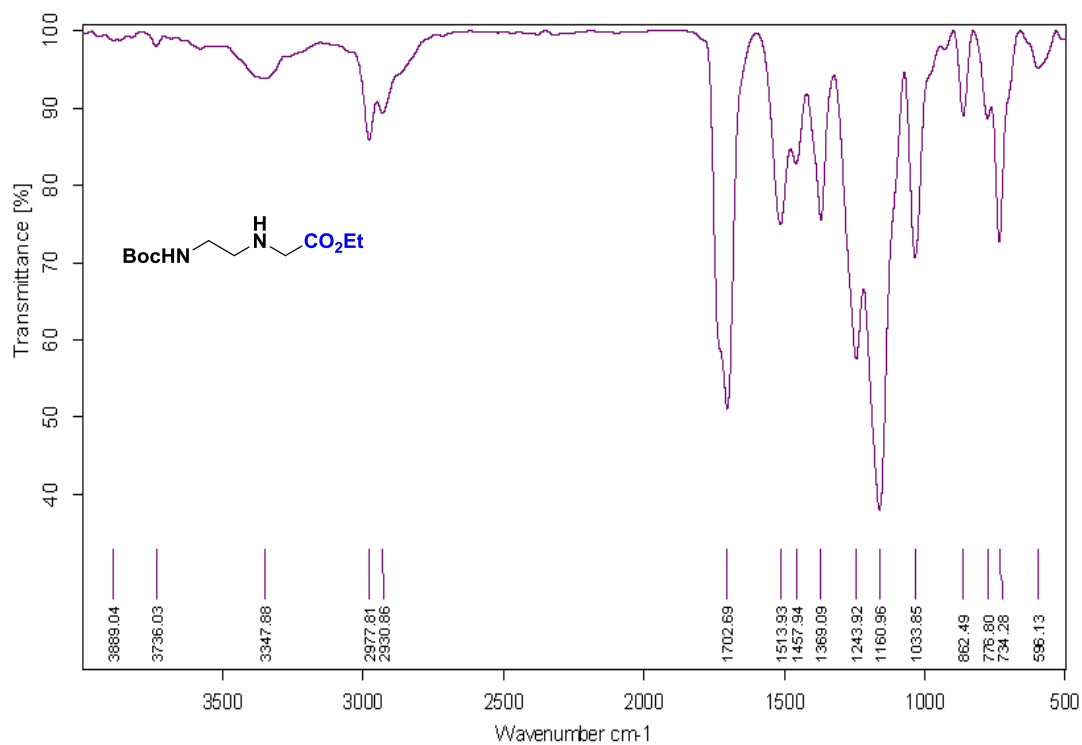
¹H NMR of Compound 31
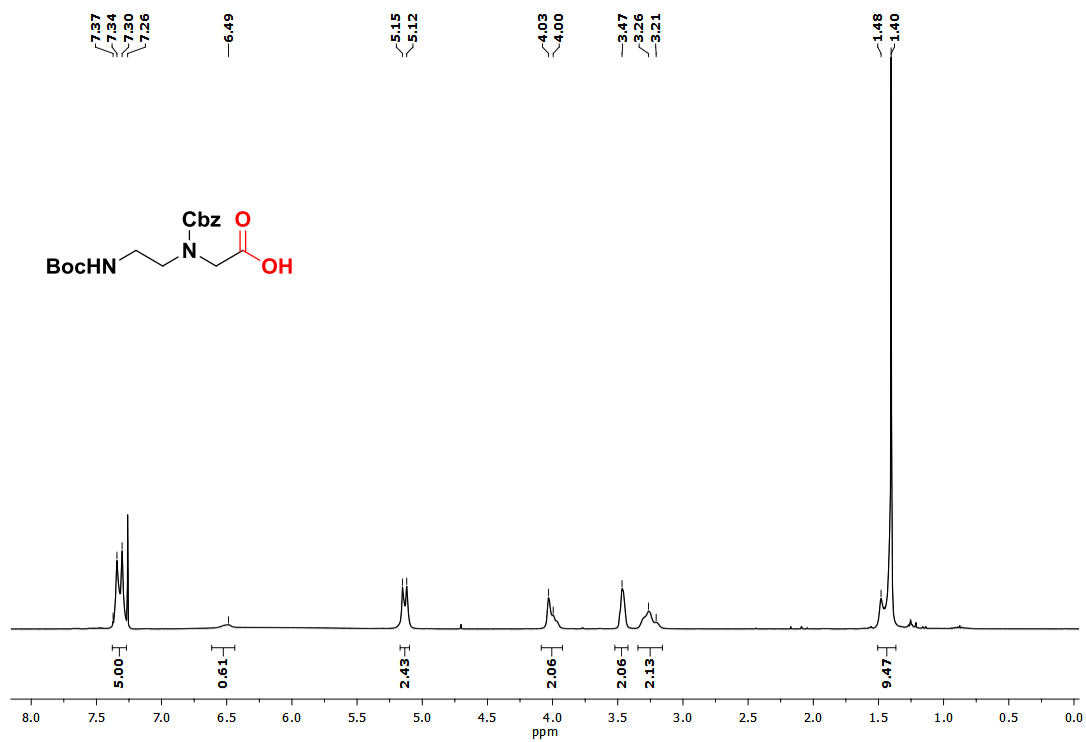
^{13}C NMR of Compound 31
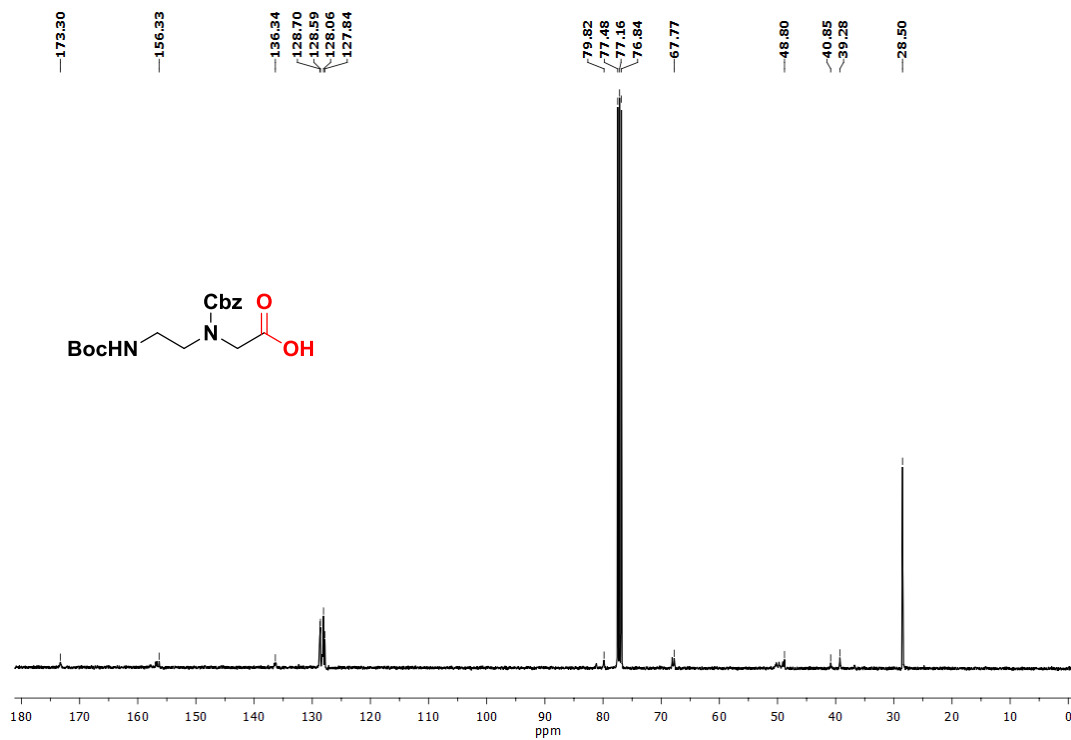
IR Spectrum of Compound 31
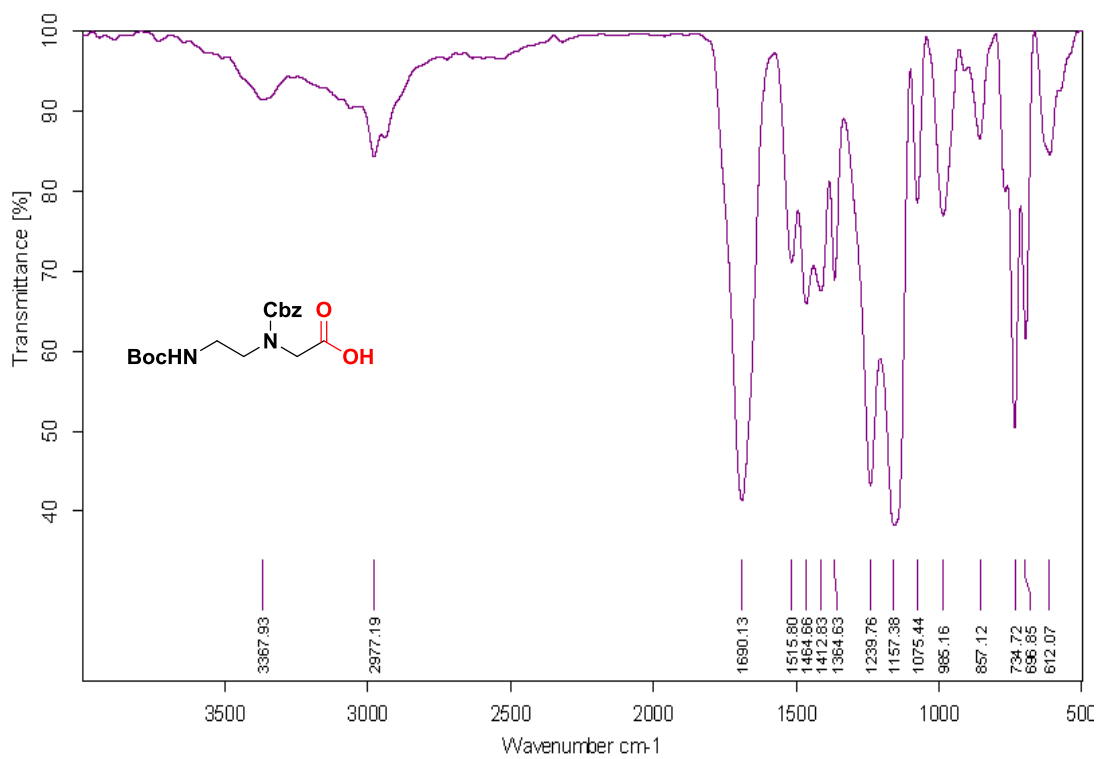
¹H NMR of Compound 34
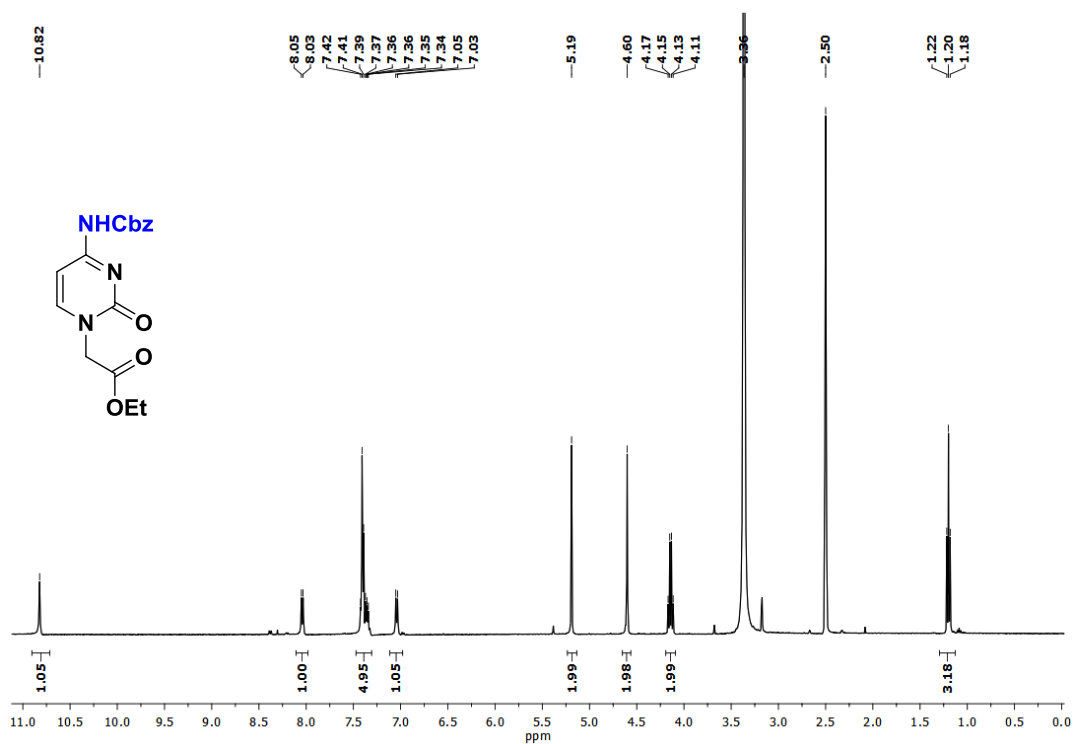
¹³C NMR of Compound 34
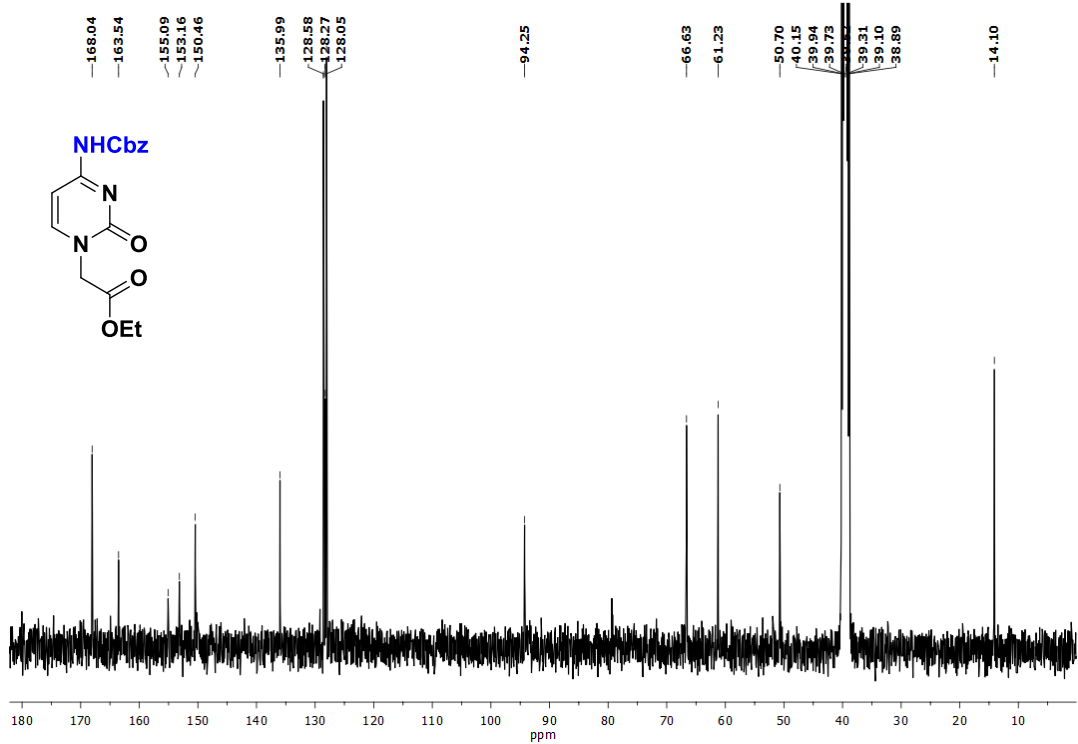
¹H NMR of Compound 35
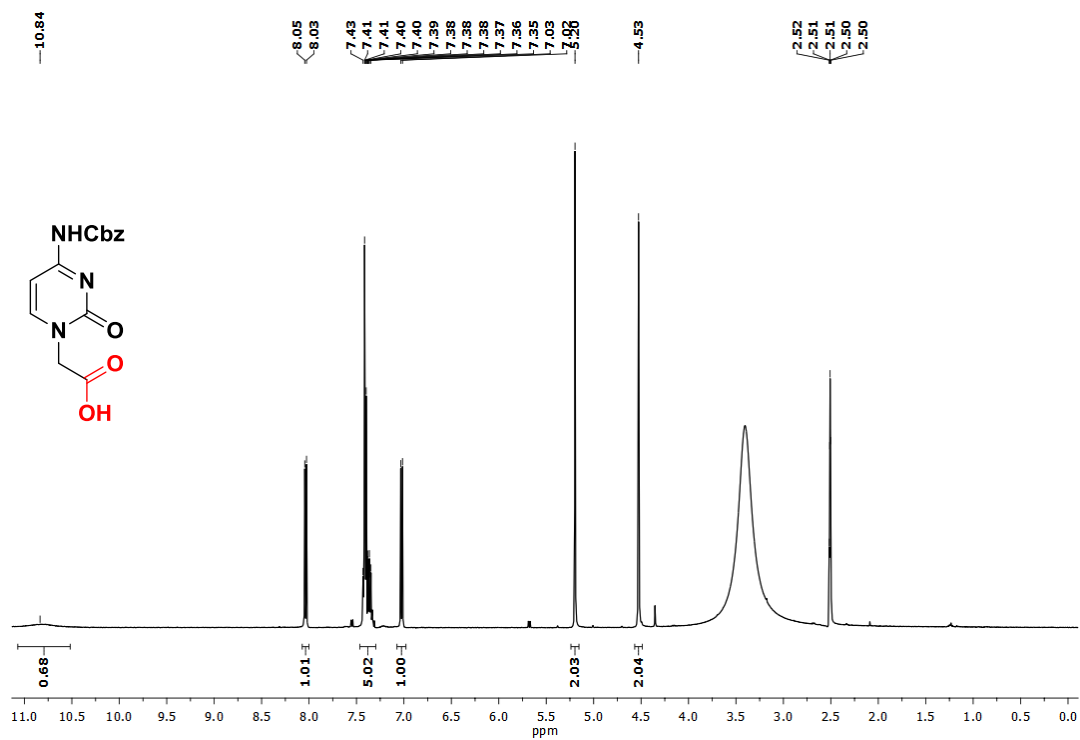
¹³C NMR of Compound 35
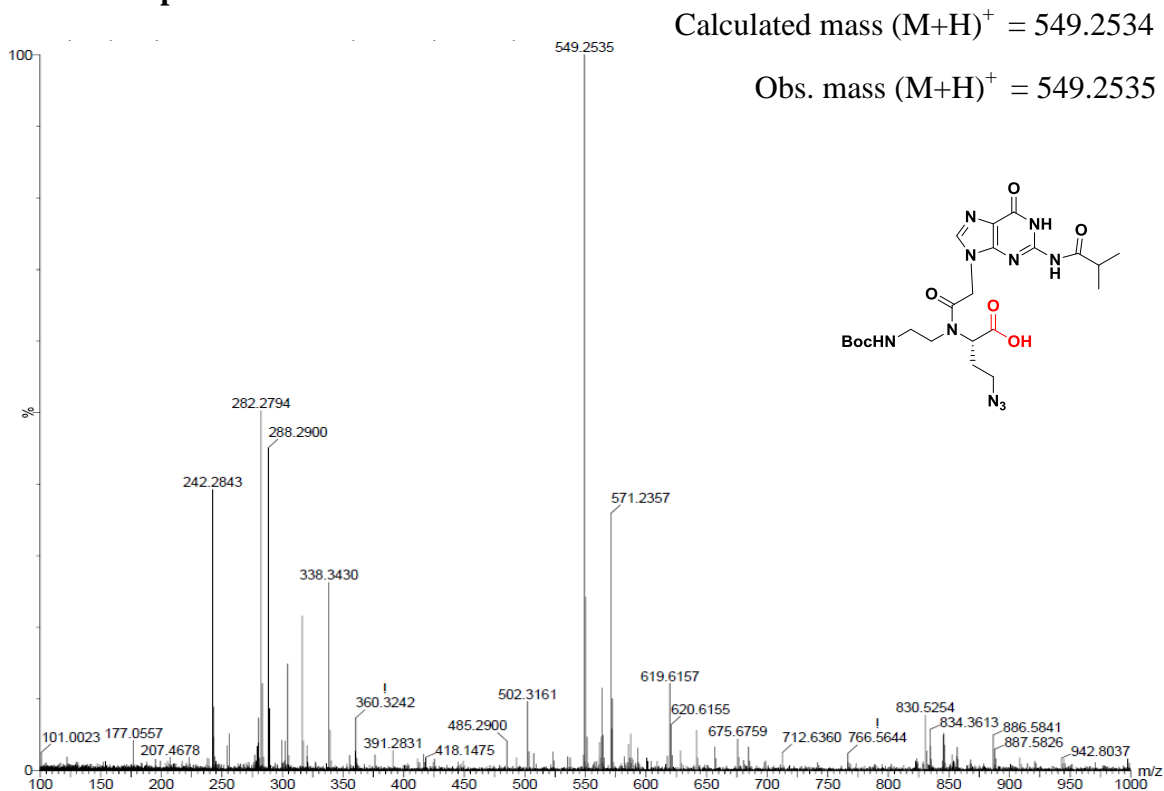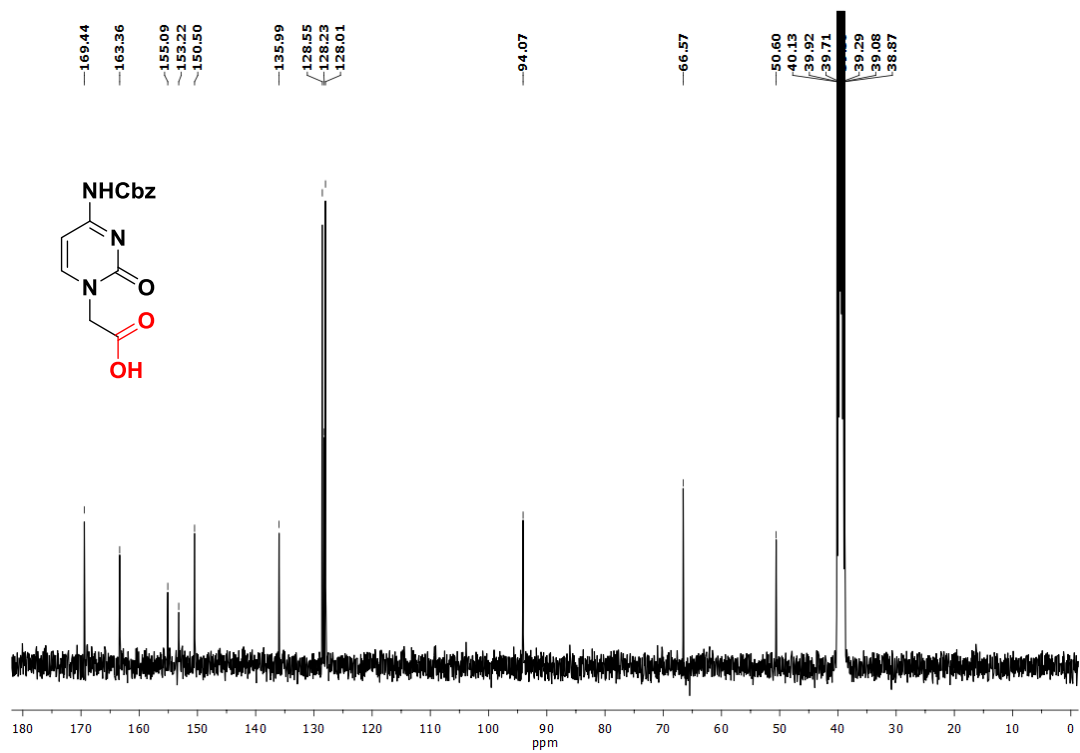
¹H NMR of Compound 37
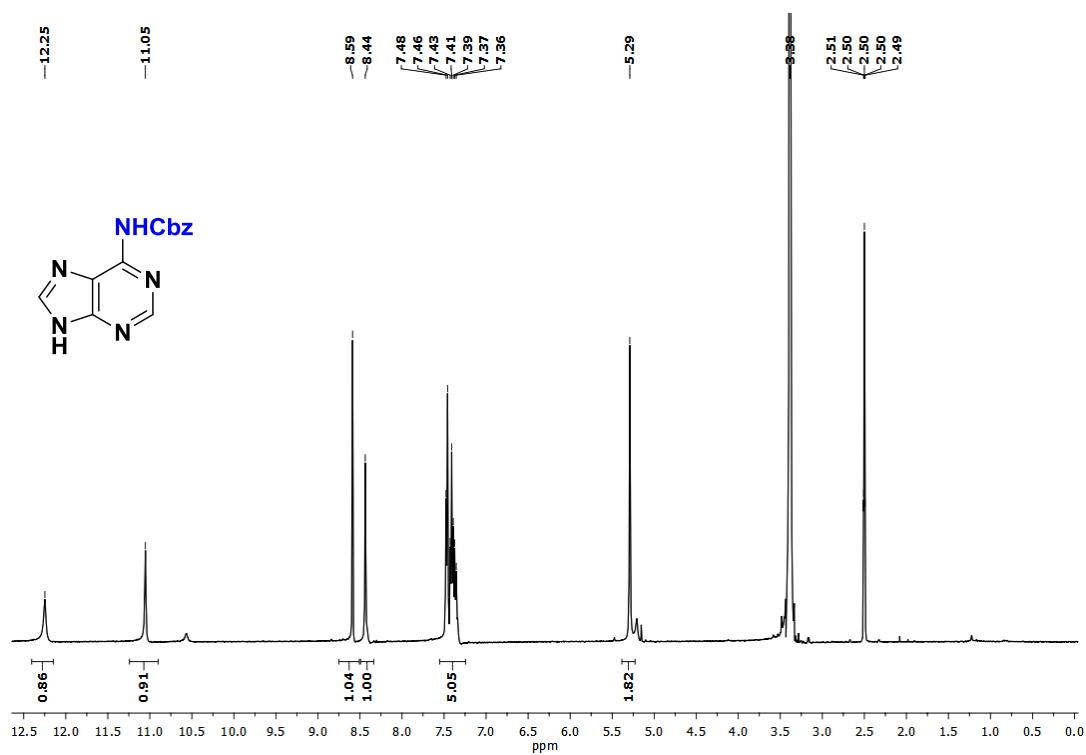
¹³C NMR of Compound 37
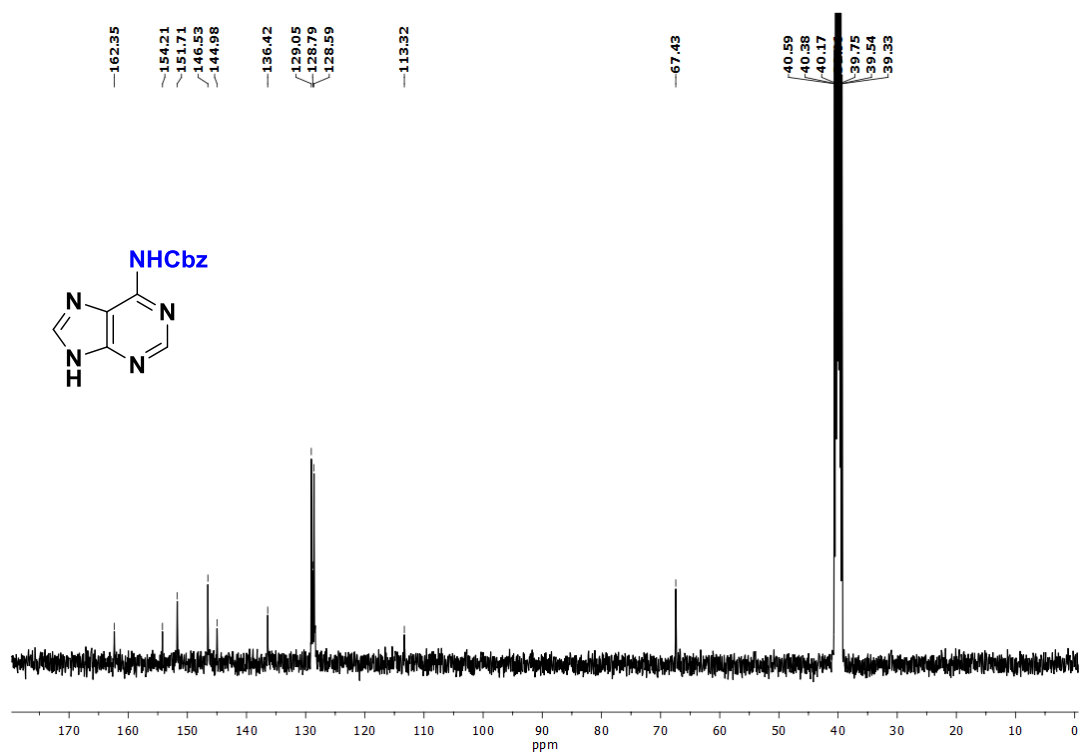
^{1}H NMR of Compound 38
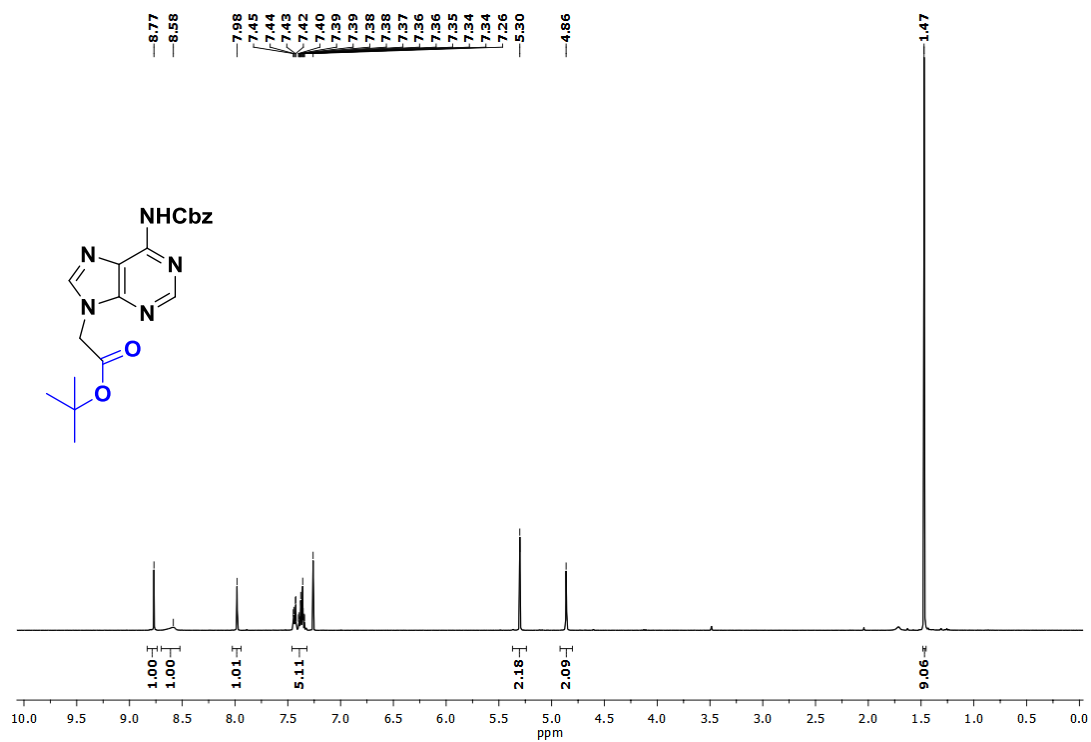
^{13}C NMR of Compound 38
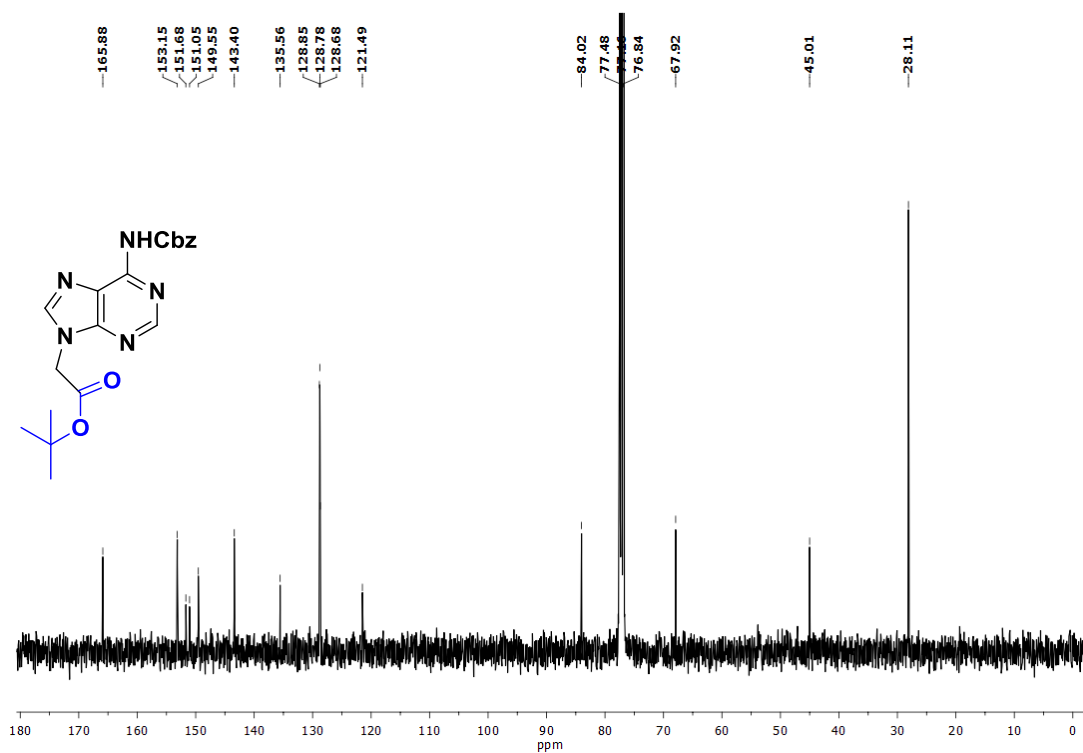
¹H NMR of Compound 39
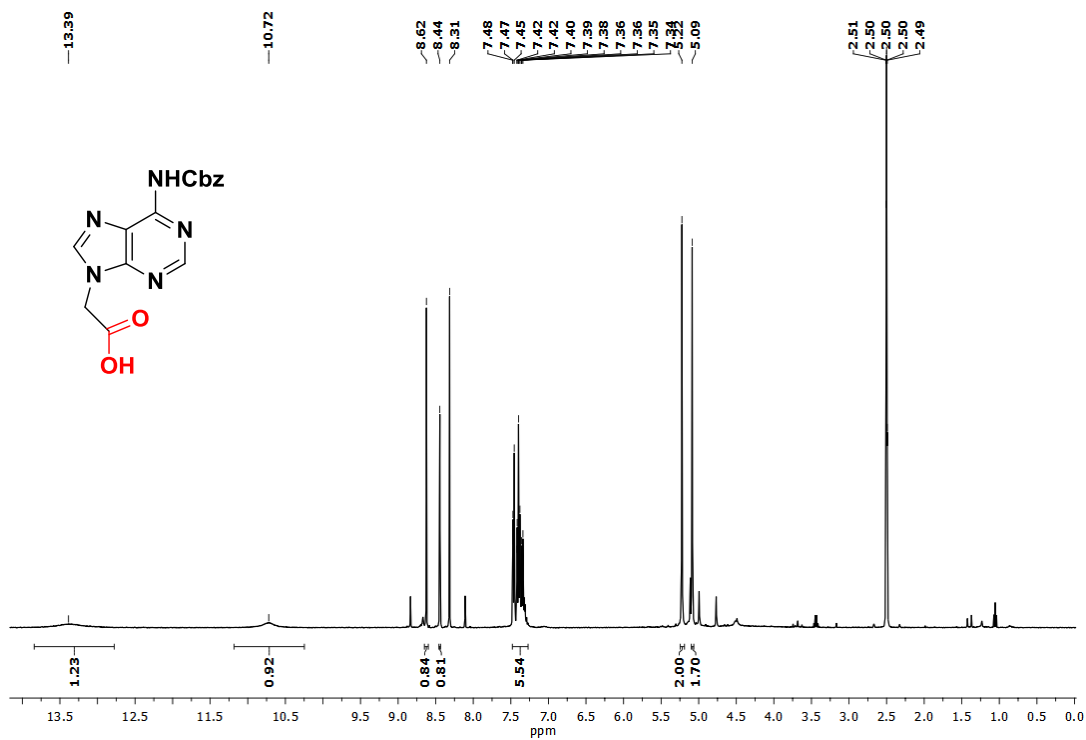
¹³C NMR of Compound 39
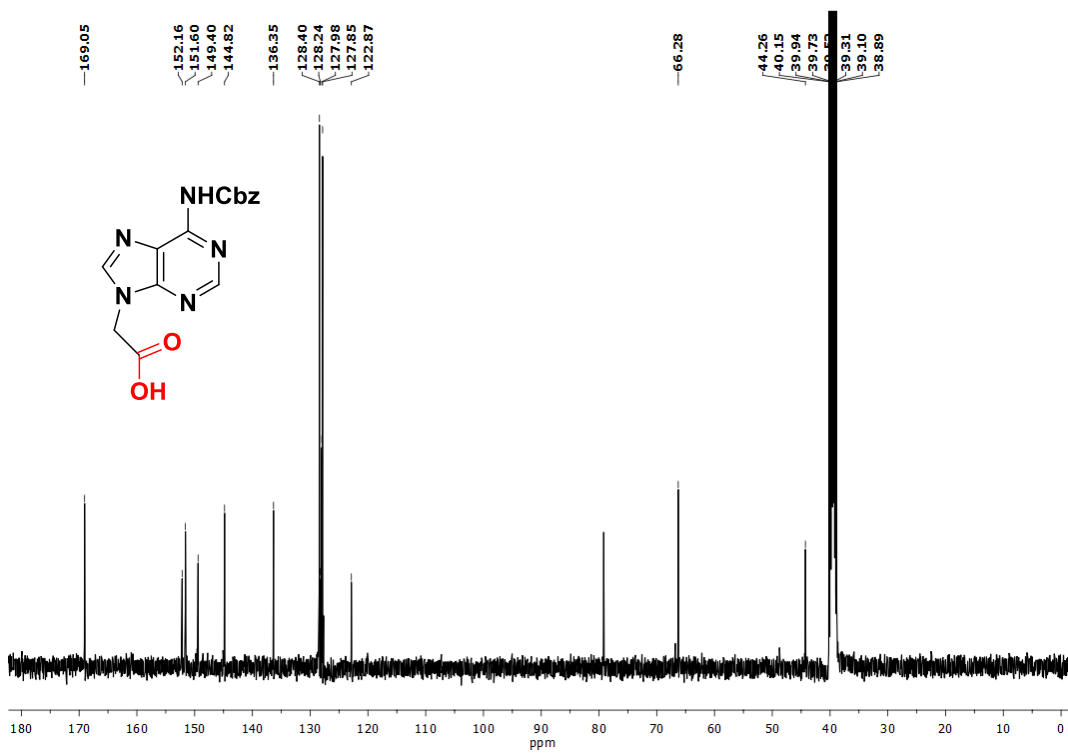
¹H NMR of Compound 42
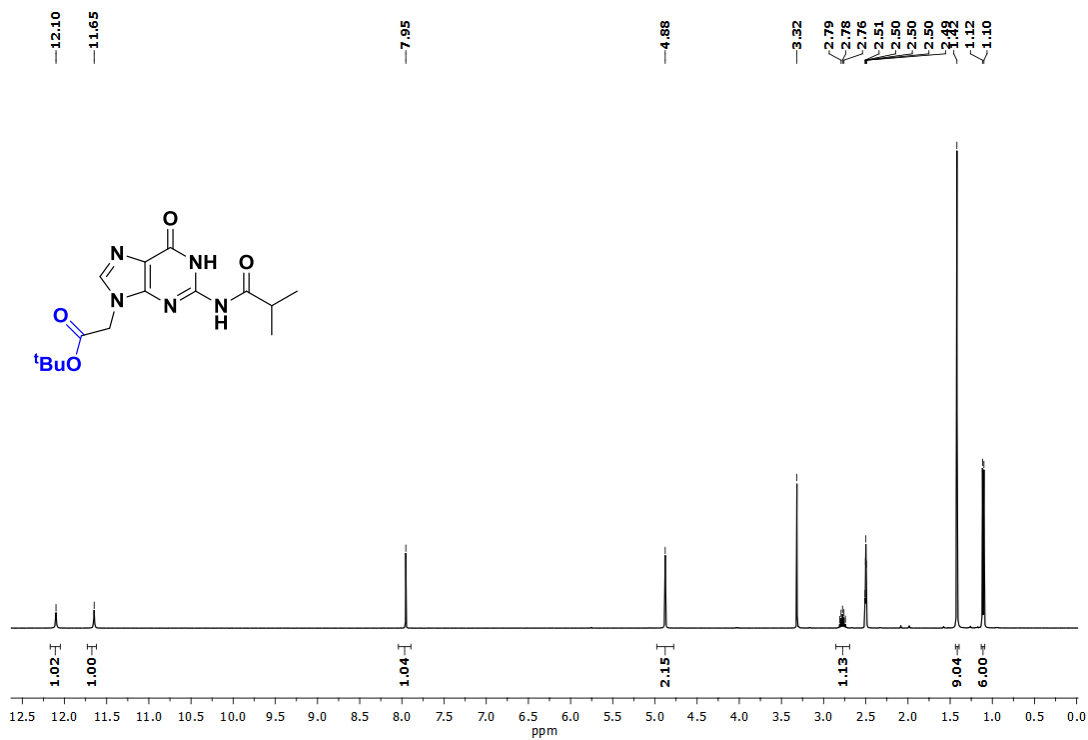
¹³C NMR of Compound 42
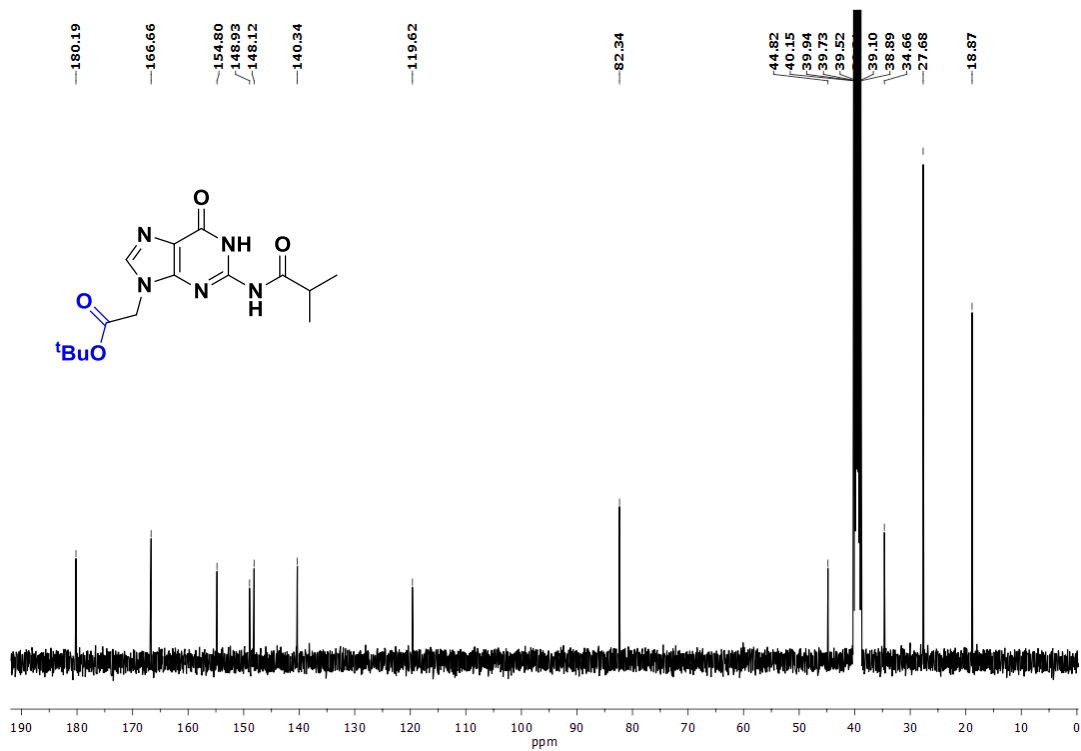
¹H NMR of Compound 43
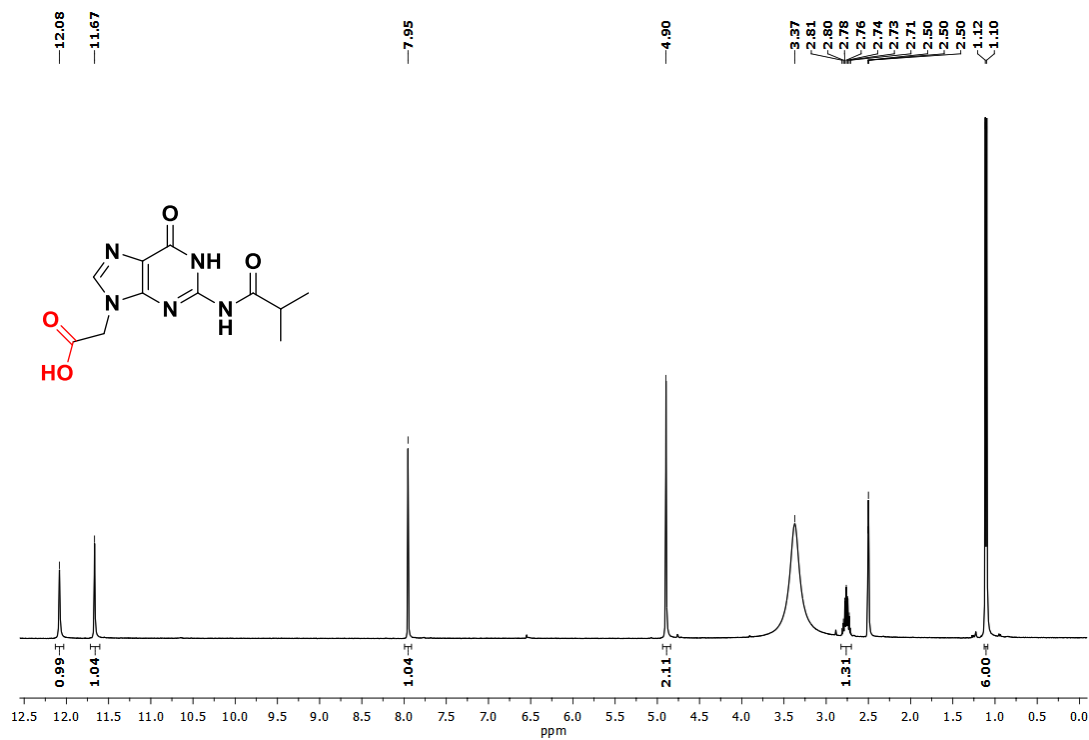
¹³C NMR of Compound 43

HRMS of Compound 43

Calculated mass $(M+H)^+ = 260.1046$

Obs. mass $(M+H)^+ = 260.1049$
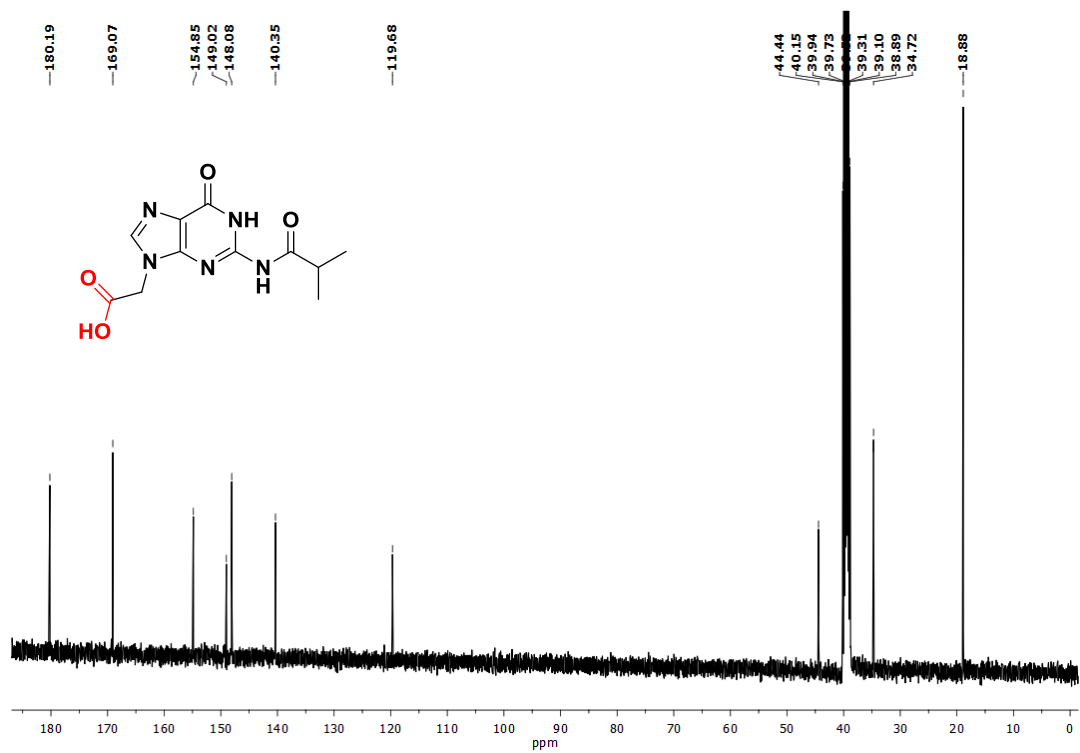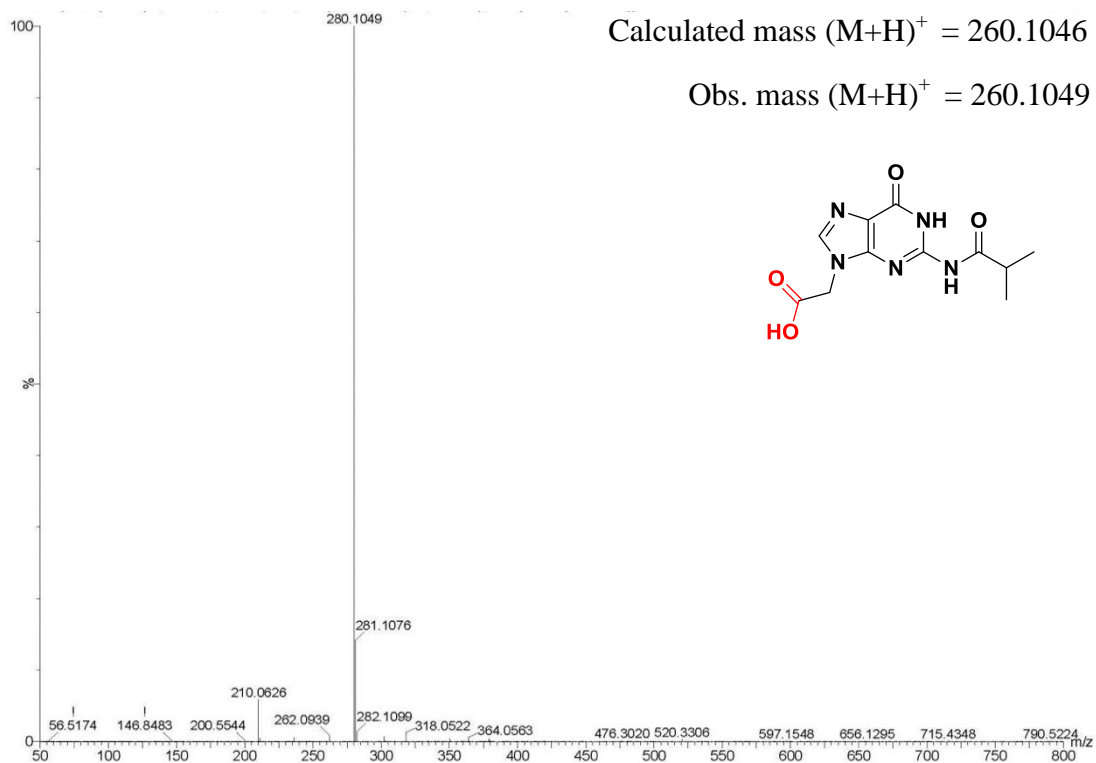
^{1}H NMR of Compound 45

^{1}H NMR of Compound 11

8.25
8.24

6.86
6.84

4.66
4.65

3.50
3.49
3.48
3.32

2.51
2.50
2.50
2.50
2.49

1.49

NBoc$_2$

1.00

1.00

2.01

0.88

17.88

9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5
ppm

^{13}C NMR of Compound 11

161.89

153.42
149.56
149.16

96.02

84.52
77.96
76.64

39.52
38.77

27.22

NBoc$_2$

180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
ppm

HRMS of Compound 11

Calculated mass $(M+Na)^+ = 372.1535$

Obs. mass $(M+Na)^+ = 372.1542$
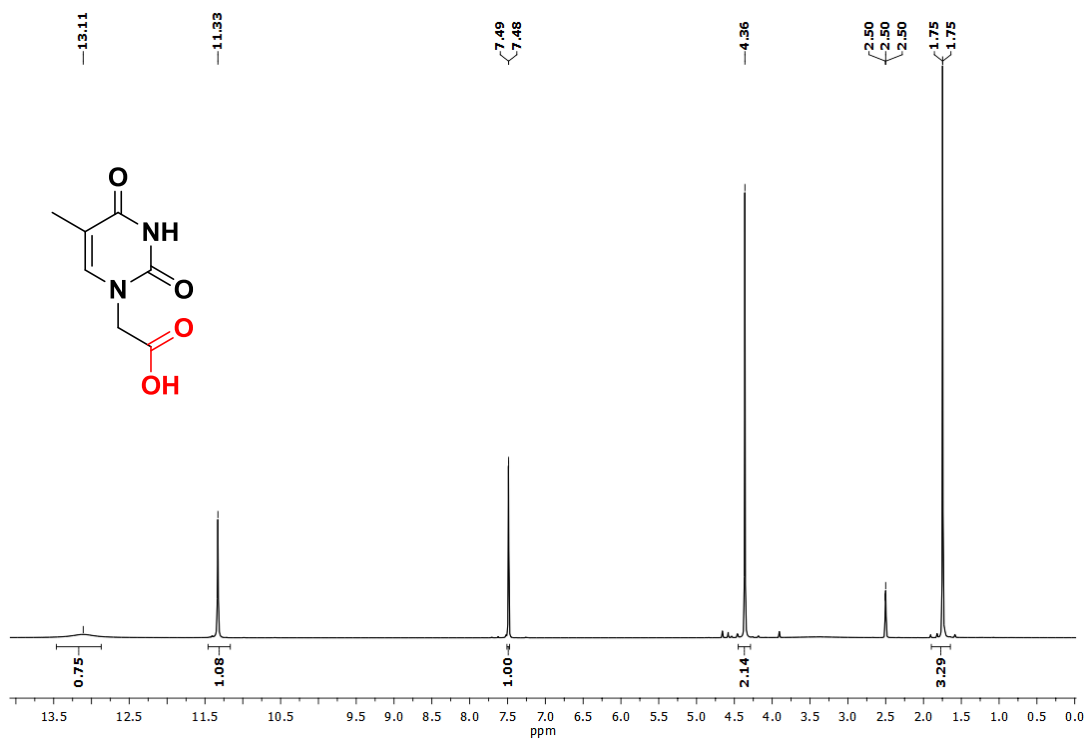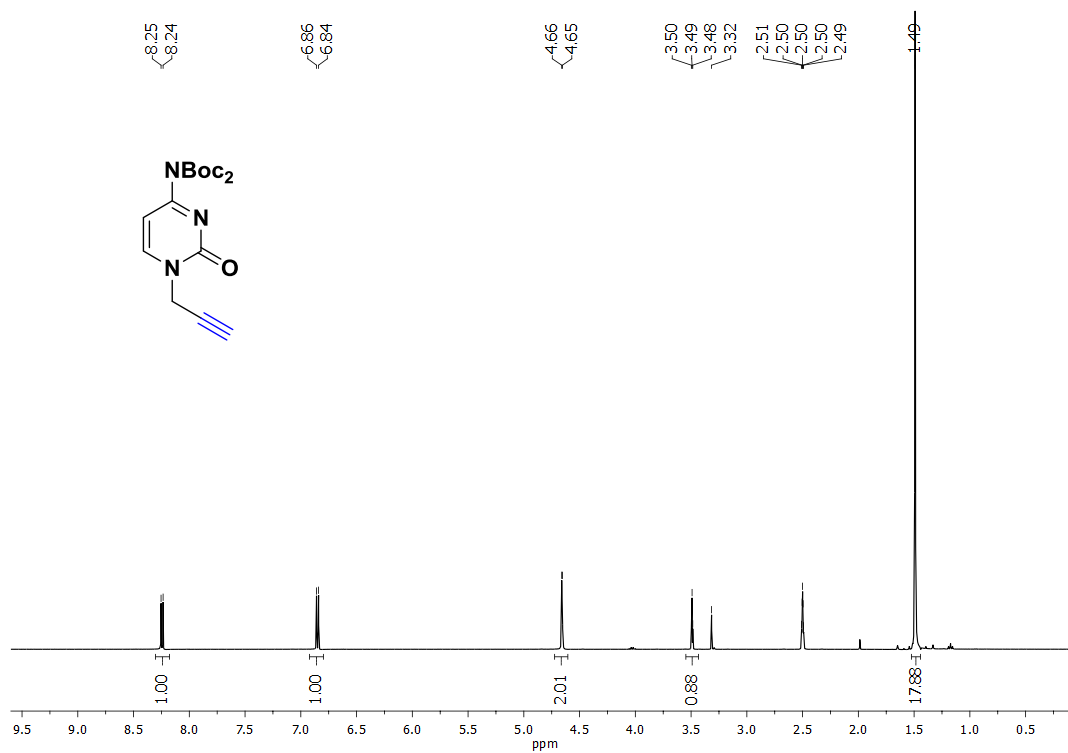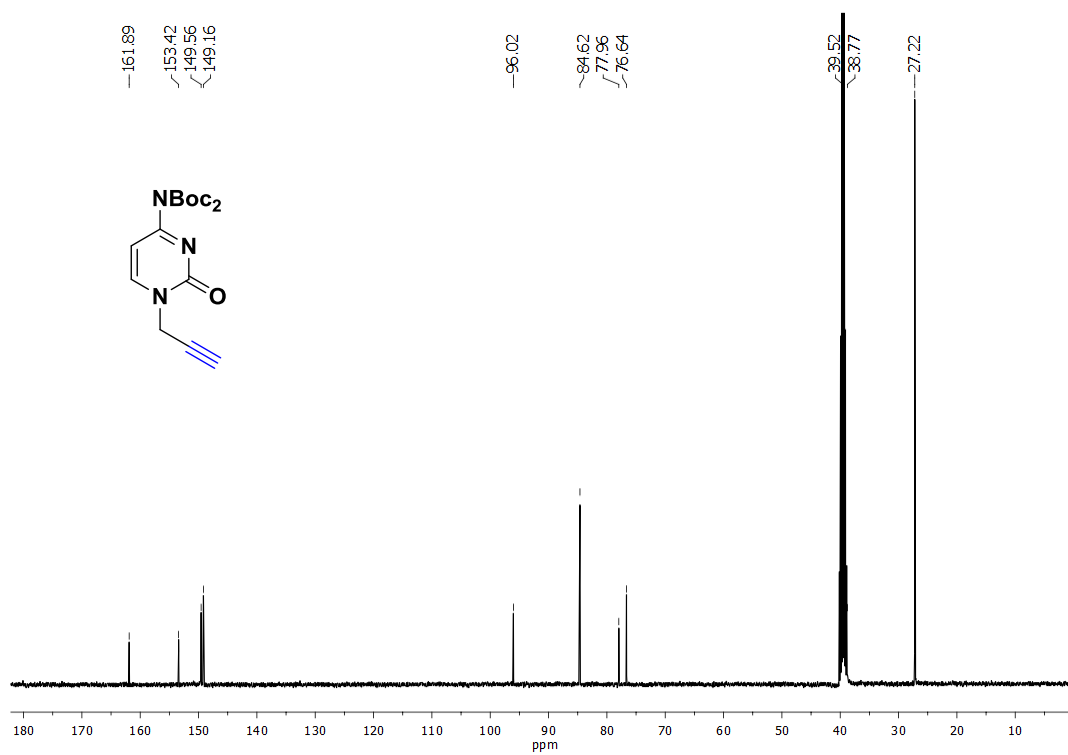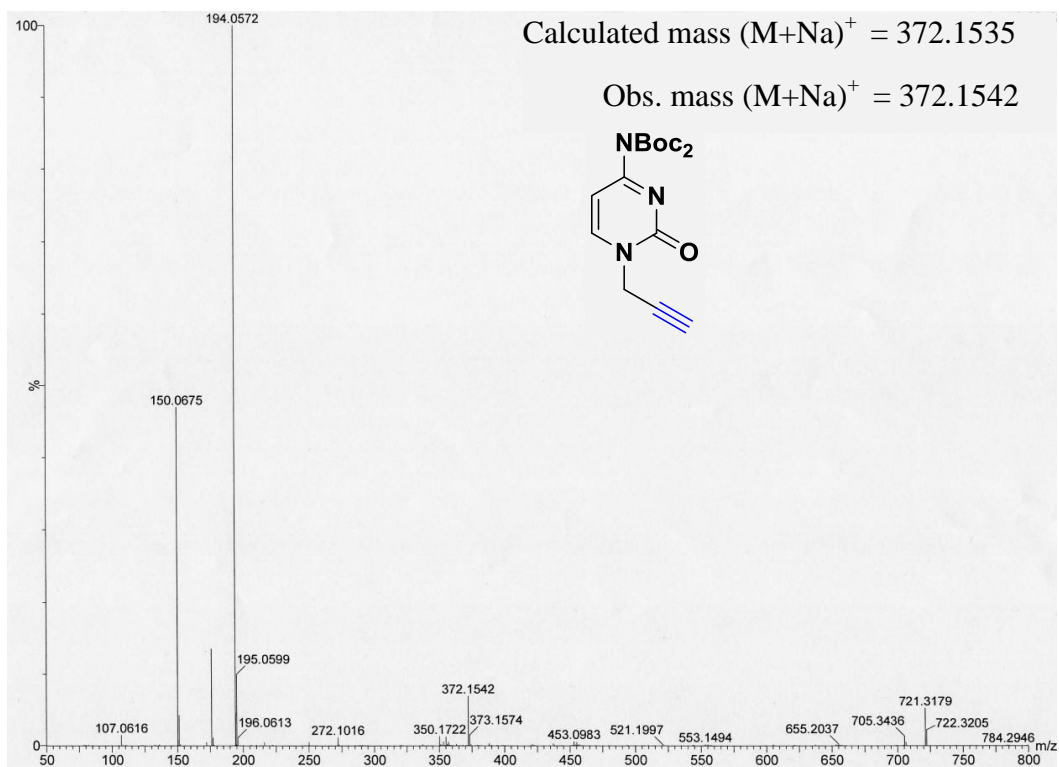
^{1}H NMR of Compound 12
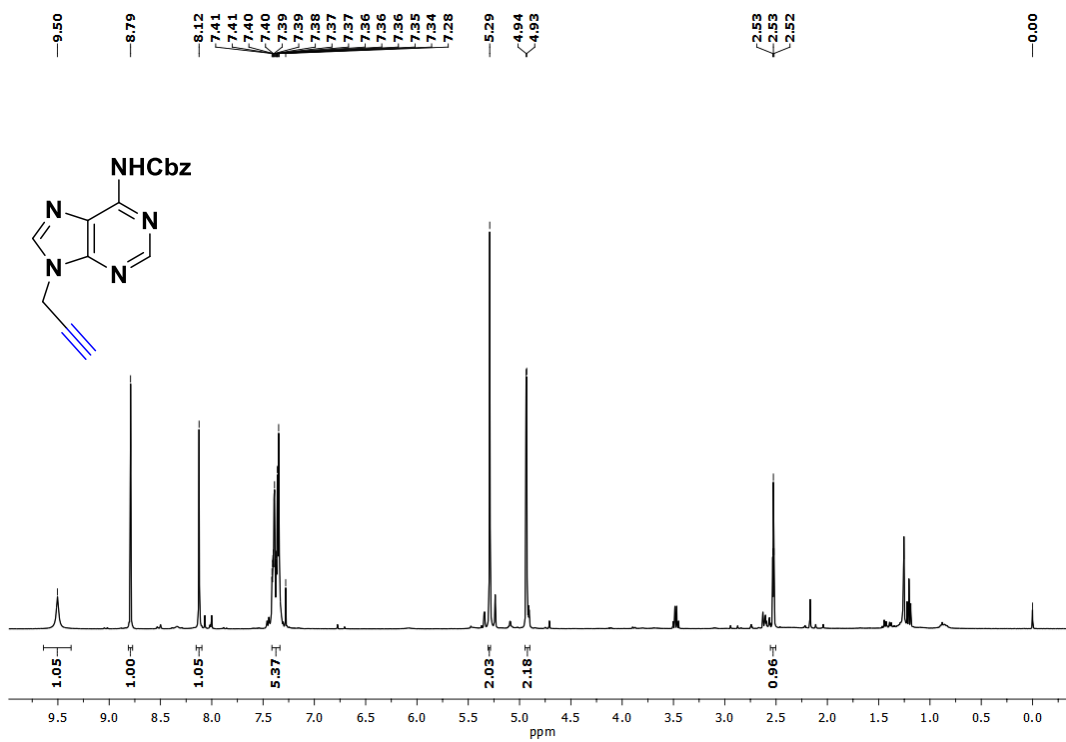
^{13}C NMR of Compound 12
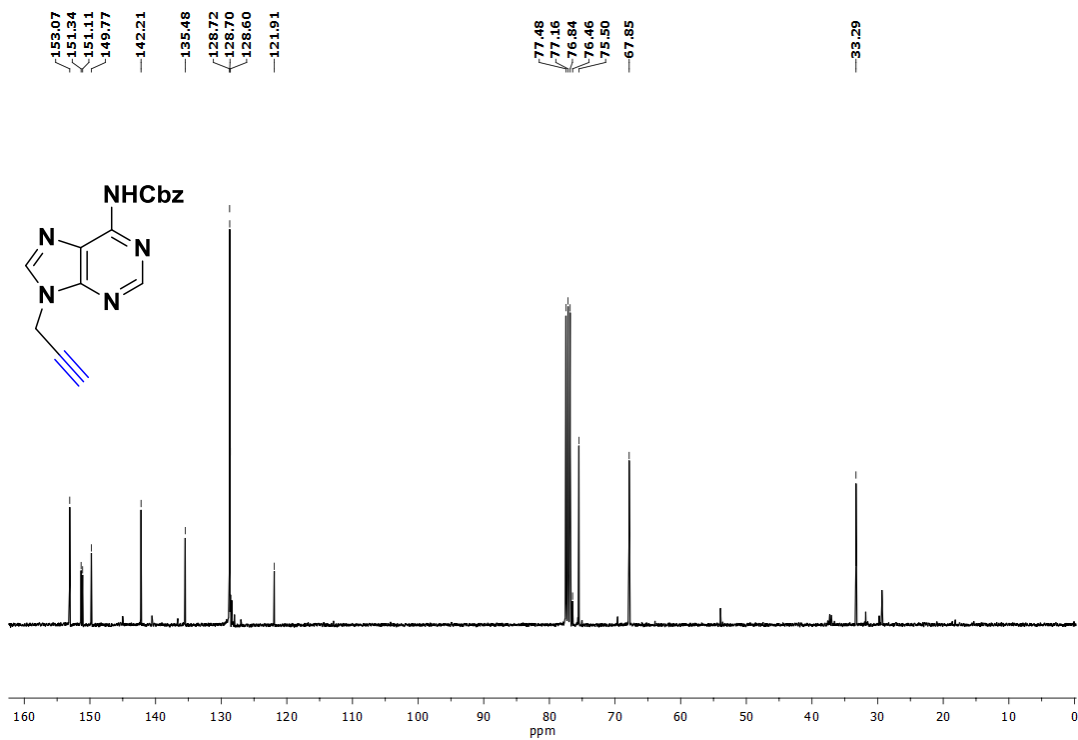
^{1}H NMR of Compound 50
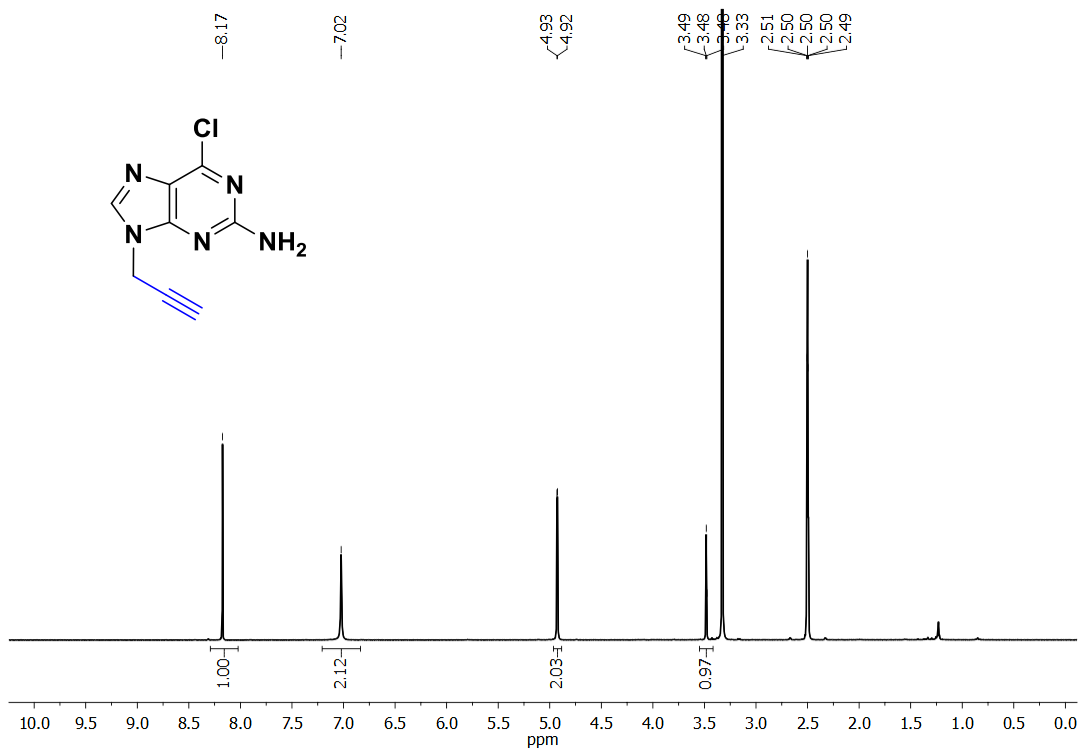
^{13}C NMR of Compound 50

HRMS of Compound 50

Calculated mass (M+H)$^+$ = 208.0390

Obs. mass (M+H)$^+$ = 208.0397
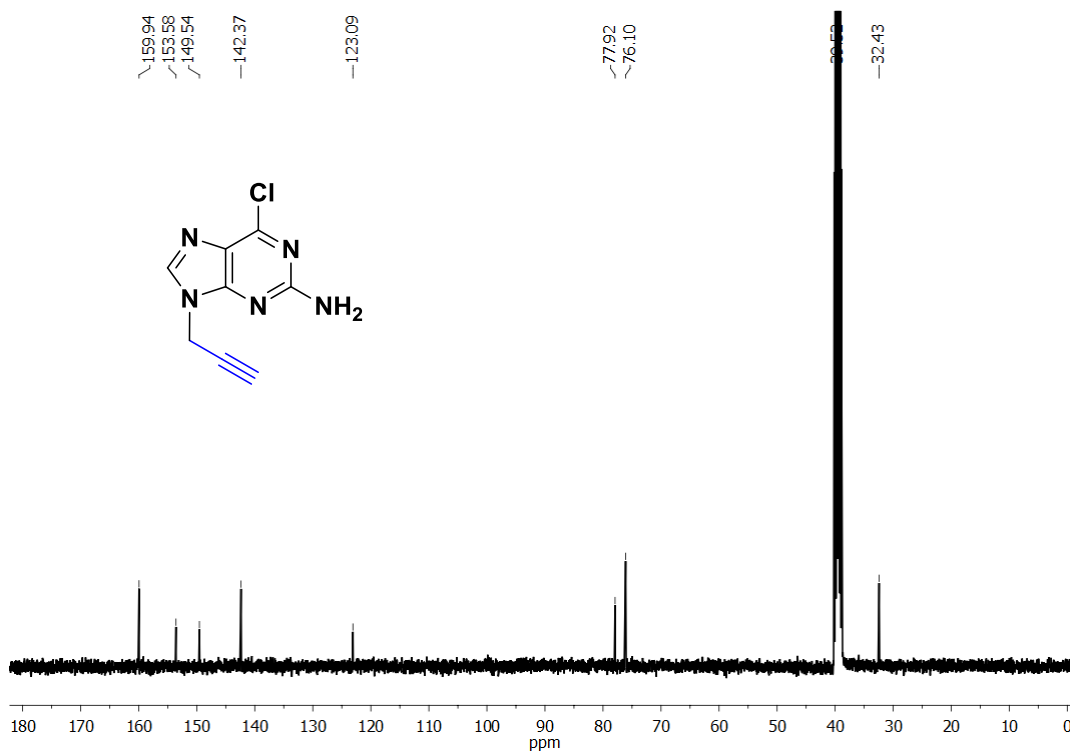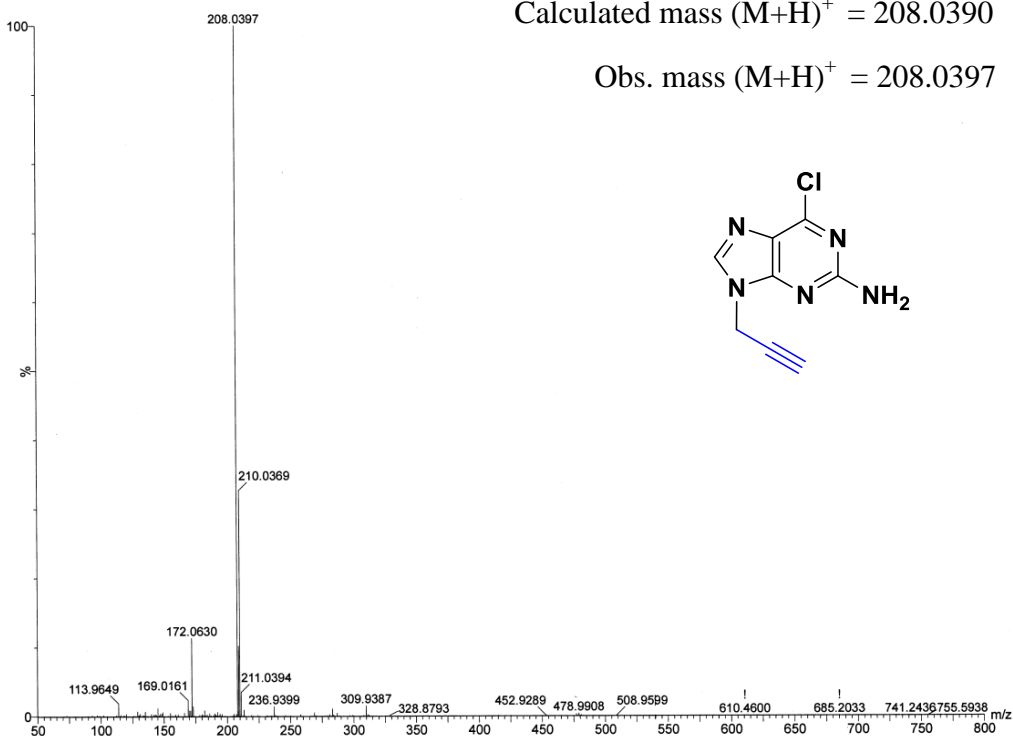
^{1}H NMR of Compound 13
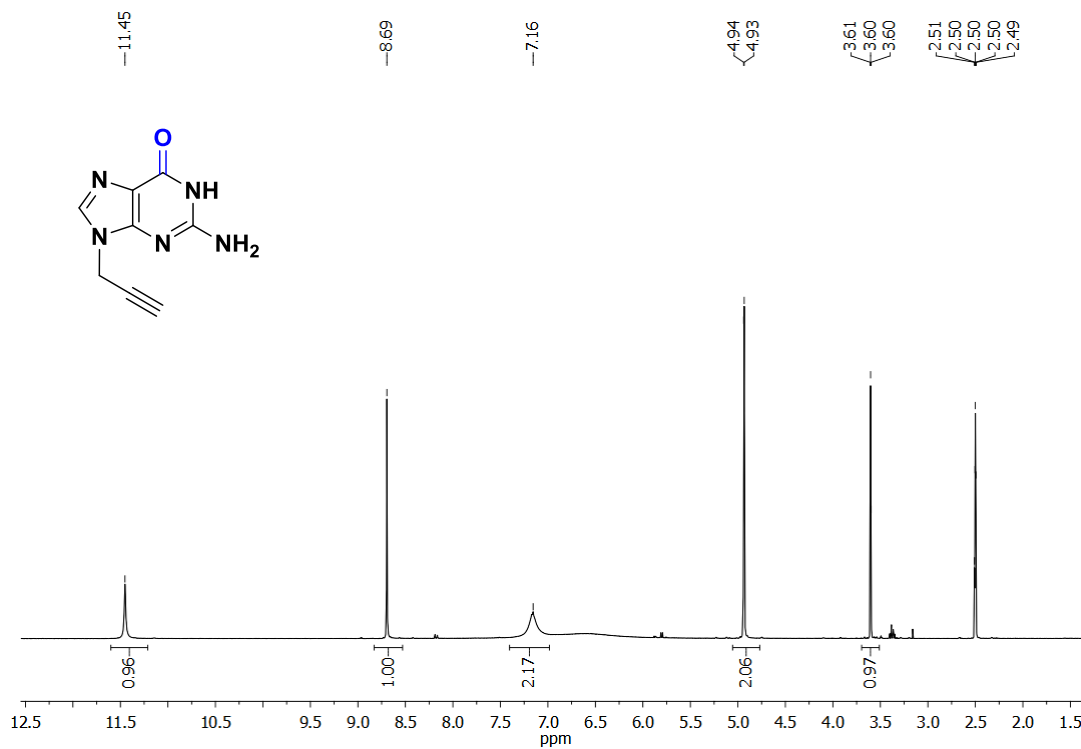
^{13}C NMR of Compound 13
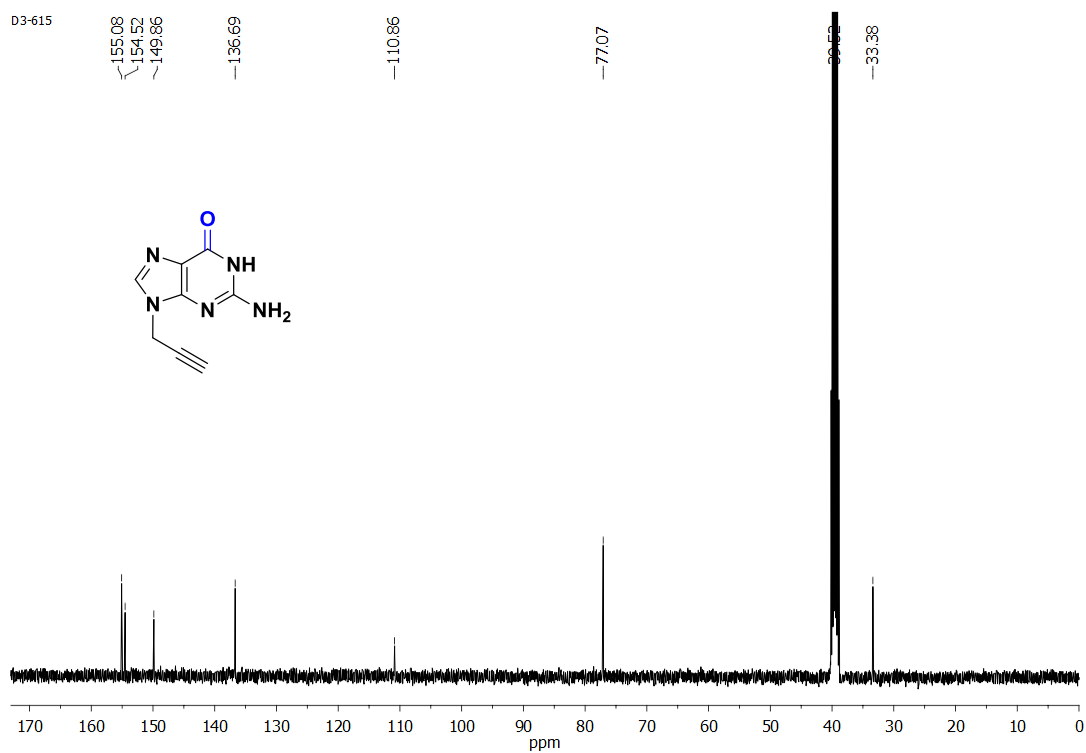
HRMS of Compound 13

Calculated mass $(M+H)^+ = 190.0729$

Obs. mass $(M+H)^+ = 190.0735$
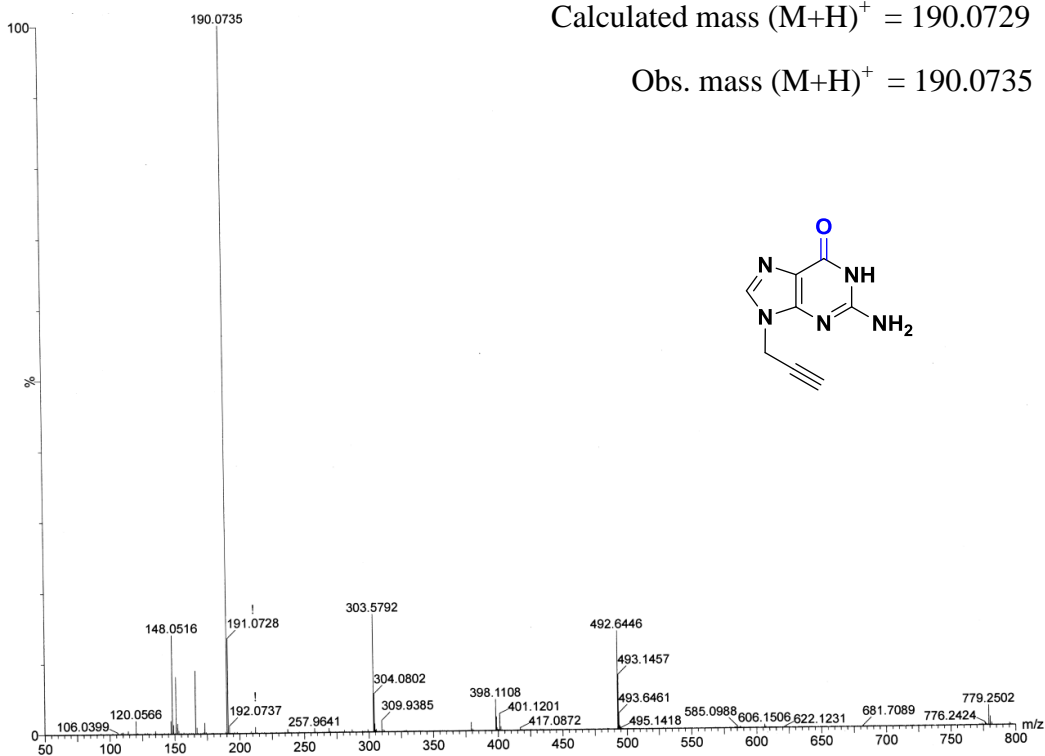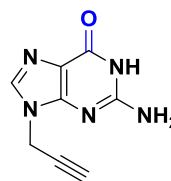
Chapter 3

Solid Phase Synthesis, Purification and Characterization of *Janus* PNA Oligomers

3.0 Aim of the present work

The specific objectives of this chapter are

➢ Incorporation of C_α-*(S-eaz) aeg* (A/T/G/C) monomers and *aeg* PNA (T/C) monomers into *Janus* PNA sequences by solid phase synthesis

➢ Click reaction of C_α-azide with base propynes on solid phase to get *Janus* PNA oligomers

➢ Cleavage of oligomers from the solid support, purification by RP-HPLC and characterization by MALDI-TOF spectrometry
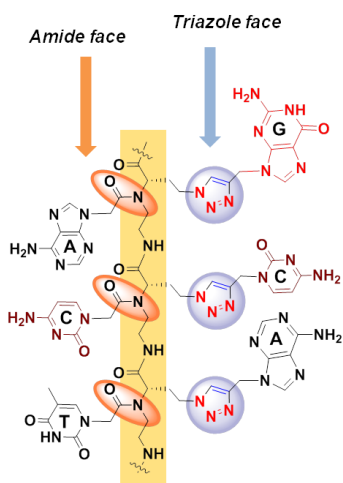
Figure 3.3. Representation of a *Janus* oligomer having an amide and a triazole face

3.1 Results and discussion

This section describes the synthesis, purification and characterization of *Janus* PNA oligomers incorporating modified (C_α-ethylazide/N^{cbz}) monomers as well as unmodified *aeg* PNA monomers into PNA sequences at desired positions. In addition to that deprotection, purification by HPLC and characterization by mass spectral data are also reported in this section.

3.1.1 Synthesis of *Janus* PNA oligomers

The synthesis of PNA oligomers was carried out manually using solid phase synthesis protocol using the Boc strategy on MBHA (4-methyl–benzhydryl amine) resin. The modified PNA monomers were incorporated at specific positions in *aeg* PNA sequence. PNA oligomerization was carried out from the C-terminus to N-terminus end using modified monomeric units with protected amino and carboxylic acid functions maintaining the orthogonality.

MBHA resin (4-methyl–benzhydryl amine resin)[1] was chosen as the solid support on which the oligomers were built and the monomers were coupled by *in situ* activation with HBTU/HOBt. In the synthesis of all oligomers, orthogonally protected (Boc/Cbz) L-lysine was selected as the *C*-terminal spacer-amino acid and it is linked to the resin through an amide bond. The amine loading on the resin was suitably reduced from 0.6 mmol/g to 0.2 mmol/g by partial acylation of amine groups using the calculated amount of acetic anhydride.[4] The free amine groups on the resin available for coupling was confirmed before starting synthesis by Kaiser's test.

The deprotection of the Boc protecting group and the completion of coupling reactions were monitored by Kaiser's test,[2,3] which is the most widely used qualitative test for the presence or absence of free amino group (deprotection/coupling). The *t*-Boc deprotection leads to a positive Kaiser's test in which the resin beads show blue color (Rheumann's purple). On the other hand, after completion of the coupling reaction, the resin beads will be colorless leading to a negative Kaiser's test. Using the standard solid phase synthesis protocol (Figure 3.4), the PNA oligomers of the desired length incorporating PNA monomers were synthesized.
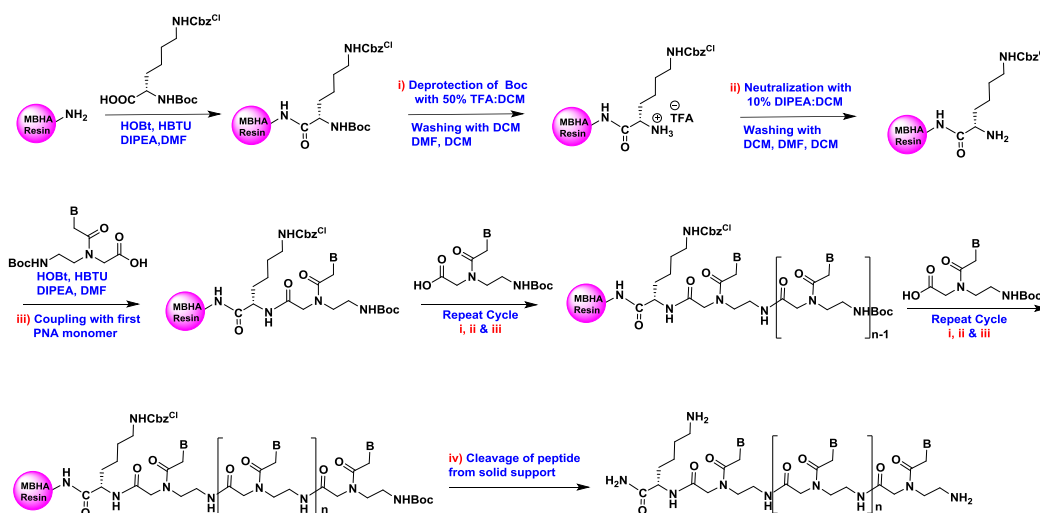
Figure 3.4 Solid phase PNA synthesis protocol by *Boc* strategy; B = Nucleobase such as Adenine (A), Guanine (G), Cytosine (C) and Thymine (T).

3.1.2 Synthesis of mixed purine-pyrimidine *Janus* PNA Oligomers

Generally, PNAs form stable duplexes of antiparallel orientation with their complementary sequences. By convention, in the antiparallel orientation of PNA:DNA duplexes, the *N*-terminus of the PNA faces the 3'- end of the DNA and C-terminus faces the 5'- end of

DNA. In parallel orientation, the *C*-terminus of PNA faces the 3'-end of the DNA and the *N*-terminus faces the 5'-end of the DNA (Figure 3.4).
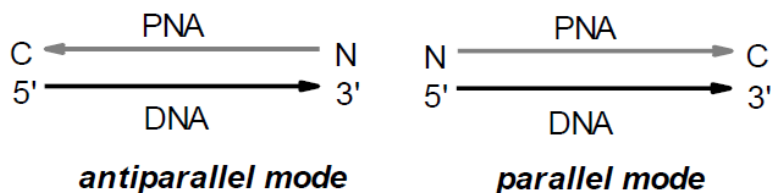
Figure 3.5 Antiparallel and parallel binding of PNA to DNA

In order to study the potential of duplex formation and the discrimination in DNA binding of the α-substituted PNAs, various kinds of *Janus* PNA oligomers were synthesized. These include -

- ➢ *Homo Janus PNA* (pyrimidine nucleobases on both faces, T on amide face and C on triazole face)
- ➢ *Chimeric Janus PNA with mixed sequence on* amide face and homo A_4 / C_5 / G_4 on triazole face
- ➢ *Hetero Janus PNA* (mixed purine-pyrimidine sequences on both amide and triazole face)
- ➢ *Self-complementary hetero Janus PNA* (mixed purine-pyrimidine sequences on both faces that are complementary to each face).

The appropriate *aeg* PNA sequences bearing A/T/C/G bases only on the amide face and PNAs having bases only on the triazole face (aminoethylglycyl C_α triazolyl PNA oligomers as controls to examine the hybridization effectiveness through double duplex formation by *Janus* PNAs through their supramolecular self-assembly.

3.2 Choice of sequence

The conformational rigidity of PNAs is enhanced by the presence of the chiral centers, that may also enhance ability of PNA strands and enrich hybridization competent conformations and, ultimately, increase the selectivity in molecular recognition. This has been demonstrated by biophysical and the crystal structure of a D-lysine-based chiral PNA-DNA duplex that has the D-lysine in place of glycine in *aeg* PNA backbone.[4-6] Taking this into consideration, PNA **9** (Figure 3.6) that contains (A/T/G/C) units on the amide face and *S*-ethylazido substitution at C_α position

of glycine was used as a primary motif for design of all *Janus* PNAs. The *homo*, *chimeric* and *hetero Janus* PNAs with mixed sequences were synthesized incorporating various kinds of modified PNA monomer units as (A/T/G/C).
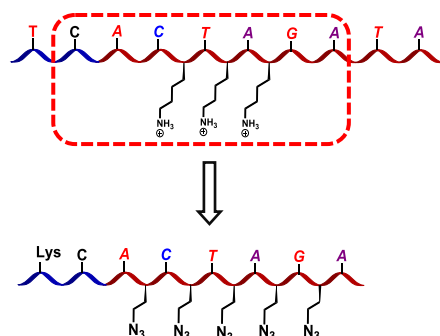
Figure 3.6 Sequence of the 7-mer PNA (PNA **9**) chosen for synthesis of Janus PNA

Based on the modifications on the two faces (amide and triazole), various kinds of *Janus*-PNA oligomers were synthesized as shown in (Figure 3.7).

i) *Homo Janus* PNA oligomers

ii) *Chimeric Janus* PNA oligomers

iii) *Hetero Janus* PNA oligomers

iv) *Self-complementary hetero Janus* PNA oligomer

iv) Aminoethyl glycyl C$_\alpha$ triazolyl PNA oligomers (*aeg-tz* PNA)
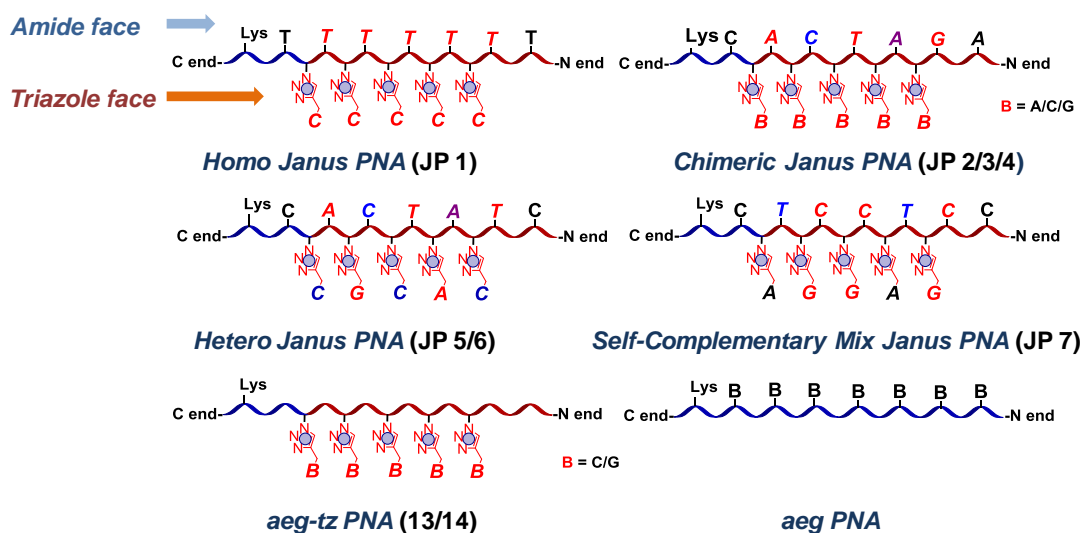
Figure 3.7 Different Types of *Janus* PNA oligomers

3.2.1 Strategy used for the synthesis of *Janus* PNA oligomers

Strategy for the synthesis of various kinds of *Janus* PNA oligomers has been described in this section.

3.2.1a Synthesis of homo Janus PNA oligomer by click chemistry on solid phase: PNA monomers were incorporated into 7-mer PNA sequence by solid phase synthesis on L-lysine derivatized MBHA functionalized resin having 0.20 mmol/g loading value. Deprotection, neutralization followed by coupling of nucleobase **4** (Figure 3.8) and repetition of the same steps gave the homo-oligomer PNA (PNA **8** or T$_\gamma$*jp-eaz$_5$*) having five ethylazide side chains. Such kind of PNA oligomer synthesis used unmodified *aeg* PNA monomer at both termini (C and N) to avoid the steric hindrance in synthesis and DNA hybridization reactions. In the subsequent step, a global click reaction of ethyl azido group on solid phase with N1-propynyl cytidine (**11**) gave solid phase bound T$_\gamma$*jp-tz-C$_5$* which was cleaved from the resin by reaction with TFA/TFMSA to yield *homo Janus* PNA oligomer (**JP 1**) T$_\gamma$*jp-tz-C$_5$* (Figure 3.8).
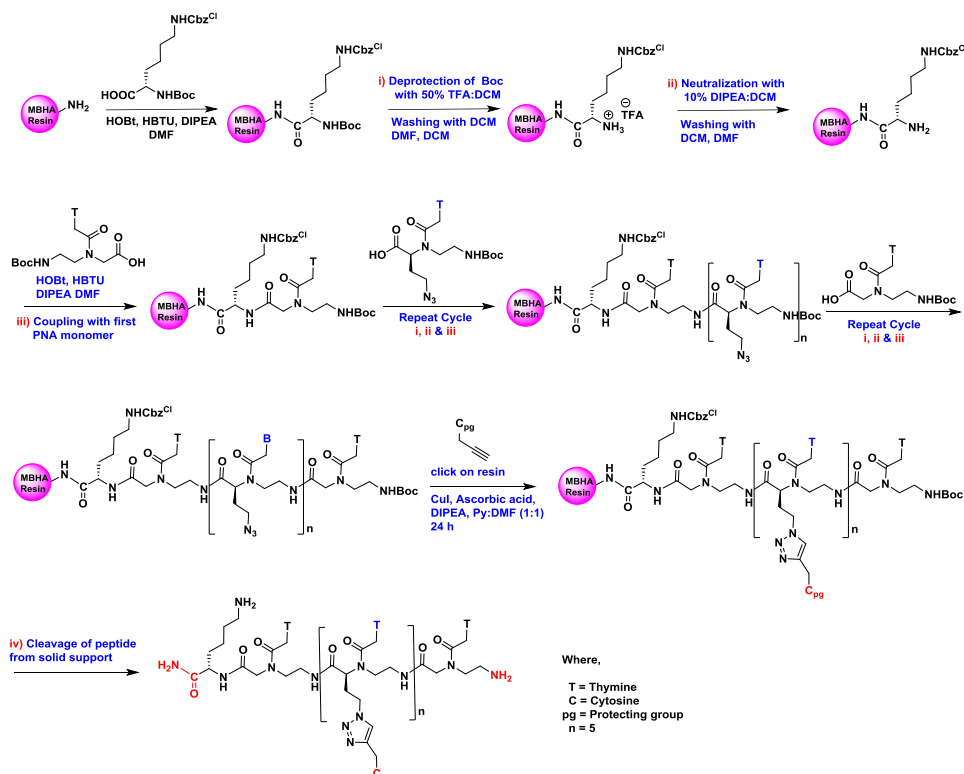
Figure 3.8 Solid phase *homo Janus* PNA (**JP 1**) synthesis protocol by *Boc* strategy
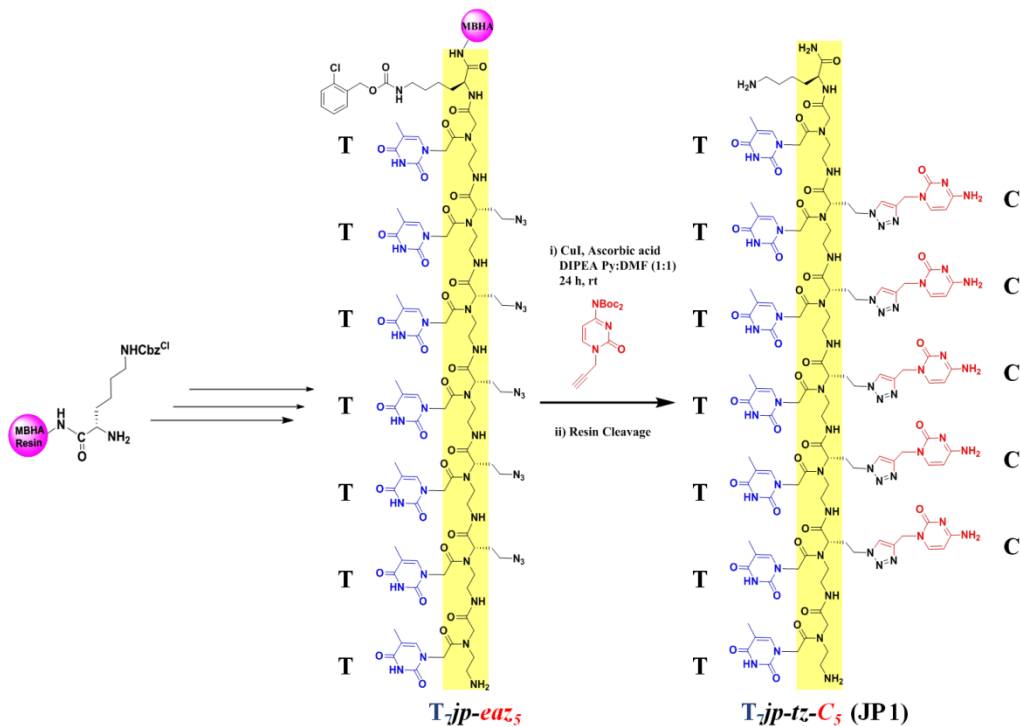
Figure 3.9 Synthetic strategy for *homo Janus PNA oligomer (***JP 1***)*

3.2.1b Synthesis of chimeric Janus PNA by click chemistry on solid phase: The same protocol as above was used for the synthesis of *chimeric Janus* PNA (**JP 2**) on solid phase, first for synthesis of *hetero* PNA (PNA **9**) having C_α-ethylazide substitution using A/G/C/T monomers (**4, 5, 6** and **7**). This yielded the solid phase linked mixed sequence PNA oligomer m*jp-eaz₅* (PNA **9**), that possess mixed purine/pyrimidines on the amide face and chimeric five $C_{\alpha-}$ethylazide side chains. To avoid the acyl migration during deprotection of isobutyryl group of guanine, N terminus was protected by an acyl group using acetic anhydride and pyridine in dry DMF.[7] This was then click-reacted individually with N1-propenyl-C (**11**),[8] N9-propenyl-A (**12**)[9] / G (**13**)[10] to get the corresponding chimeric PNAs m*jp-tz-C₅* (**JP 2**), m*jp-tz-A₄* (**JP 3**), and m*jp-tz-G₄* (**JP 4**), (Figure 3.10).
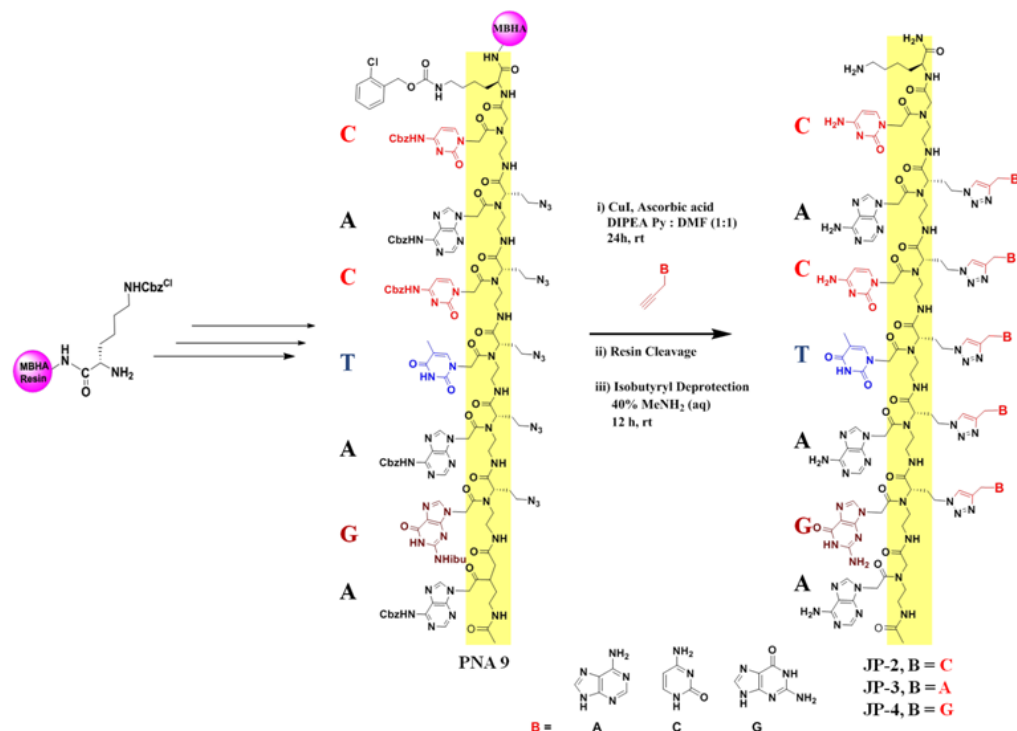
Figure 3.10 Synthetic strategy of *chimeric Janus PNA oligomers (***JP 2***, **JP 3** *and* **JP 4***)*

3.2.1c Synthesis of hetero Janus PNA with mixed sequence on both amide and triazole face:
The *hetero Janus* PNAs that possess mixed purine/pyrimidine bases on both the amide and triazole faces were synthesized by the same solid phase protocol as above, using MBHA resin and *S*-ethylazide PNA (A/T/G/C) monomers. As compared to strategies shown in Figures 3.9 and 3.10, click reactions were done after each monomer coupling steps (Figure 3.11), using appropriate propyne bases (A/G) to assemble the triazole face sequence. In between some steps, click reaction was avoided, allowing some free azide groups to be present in the sequence. At the end of synthesis these azide groups were reacted with N1-propenyl cytosine[21] (**11**) to install C base at these sites (Figure 3.12). The resultant lead to mixed sequences on the triazole face as well as the amide face. In order to develop easier and flexible protocols for such *hetero Janus* PNAs having mixed sequence on both sides, *Janus* monomer **10** having pre-installed base T on amide face and A on triazole face was synthesized and used for coupling on solid phase. Such an approach avoids click reaction on the solid phase and the availability of different combinations of *Janus* monomers with two bases and opportunity for the synthesis of *hetero Janus* PNA oligomers (**JP 5** and **JP 6**) of desired mixed sequence on both faces. The synthesized *hetero Janus* PNA was cleaved from resin purified by HPLC and characterized by mass spectral data.

To ensure successful and efficient coupling at all steps, a small amount of the resin was cleaved at intermediate stages of synthesis at difficult steps of coupling and the intermediate products were characterized by mass spectral data.
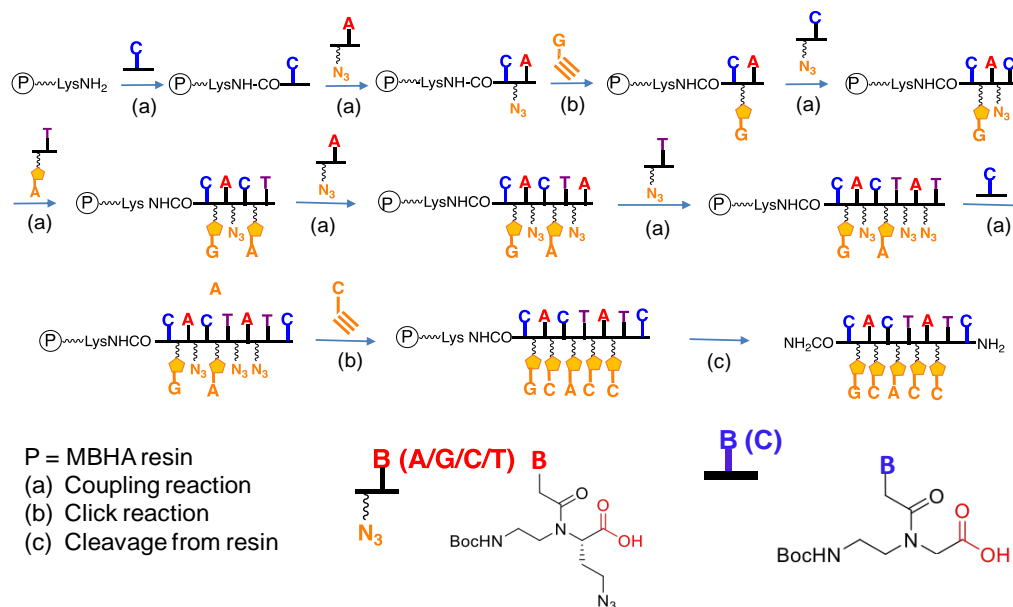
Figure 3.11 Synthesis strategy of m*jp-tz₅-CCACG* oligomer (**JP 6**) on solid support
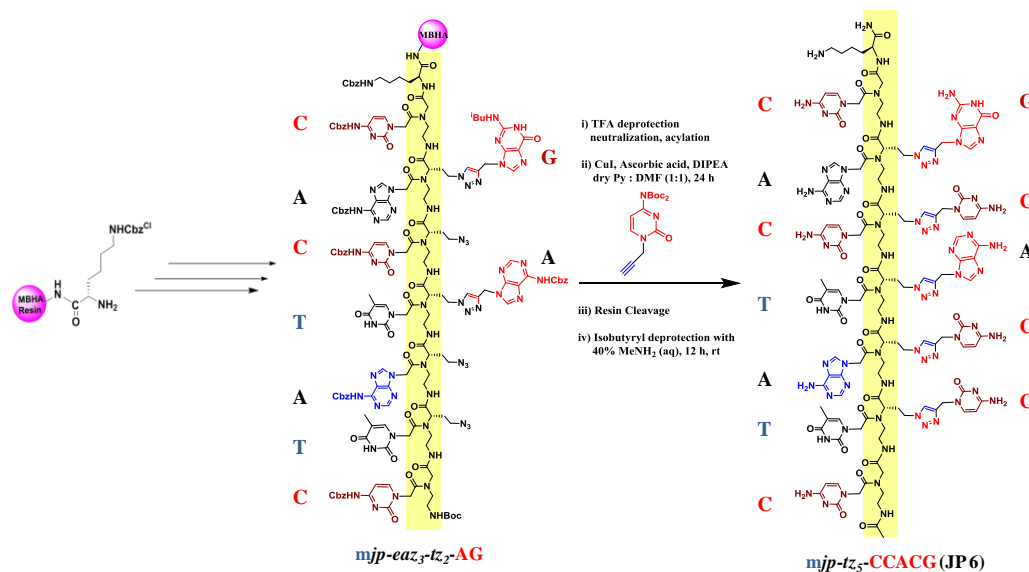
Figure 3.12 Synthetic strategy of m*jp-tz₅-CCACG* oligomer (**JP 6**)

3.2.1d *Synthesis of self-complementary hetero Janus PNA by click chemistry on solid phase:* Synthesis of *Janus* PNA oligomer (**JP 7**) that possesses mixed purine and pyrimidine sequence on amide face and triazole faces that are complementary to each other (*self-complementary*

hetero Janus PNA) (Figure 3.13) was synthesized by the same protocol as depicted in Figure 3.13. Such *self-complementary Janus* PNAs would have the potential to form supramolecular assemblies through self-hybridization. The synthesized PNA **JP 7** was purified by HPLC and characterized by mass spectral data.
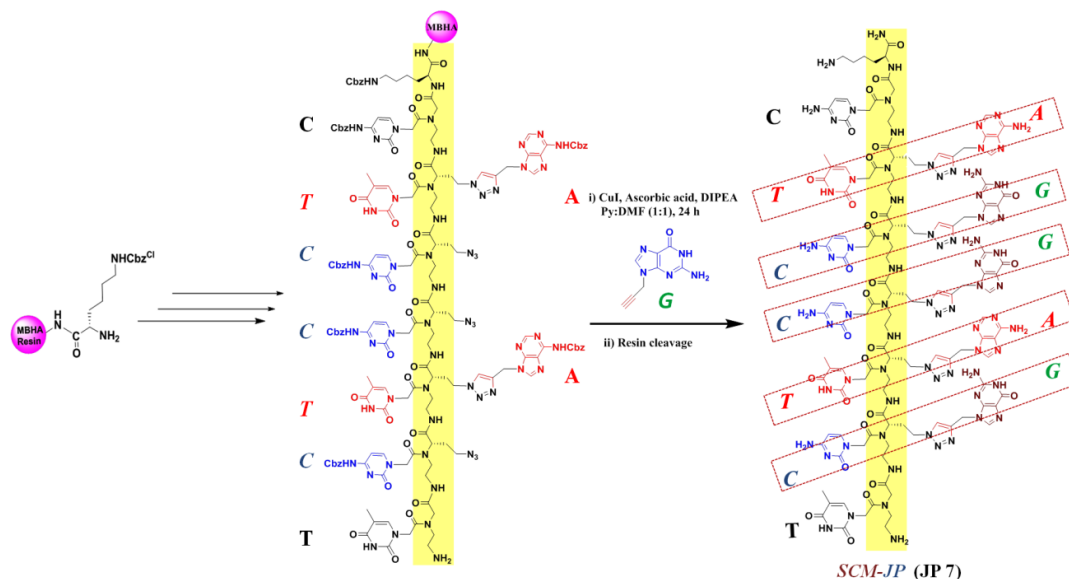
Figure 3.13 Synthesis strategy of *self-complementary mix Janus* PNA (**JP 7**) oligomers

3.2.1e *Synthesis of PNA analogue with nucleobase on only aminoethyl glycyl triazole face:* For studying hybridization exclusively from triazole face, synthesis of aminoethylglycine backbone oligomers having only C_α substituted triazole face bases (*aeg-tz* PNA), but no bases on amide face were synthesized (Figure 3.14). This involved the use of modified monomer α-(*S-eaz*)]-N^{Cbz}-*aeg* (**8**) and click-reacted individually on resin with N1-propenyl-C (**11**), N9-propenyl- G (**12**) to get the corresponding triazolyl derived PNAs p7-*tz-C$_5$* (**PNA 13**) and p7-*tz-G$_5$* (**PNA 14**) with a secondary NH group present on the backbone. After solid phase assembly, these were purified and characterized by mass spectral data in the same way as described for other *Janus* PNA oligomers.
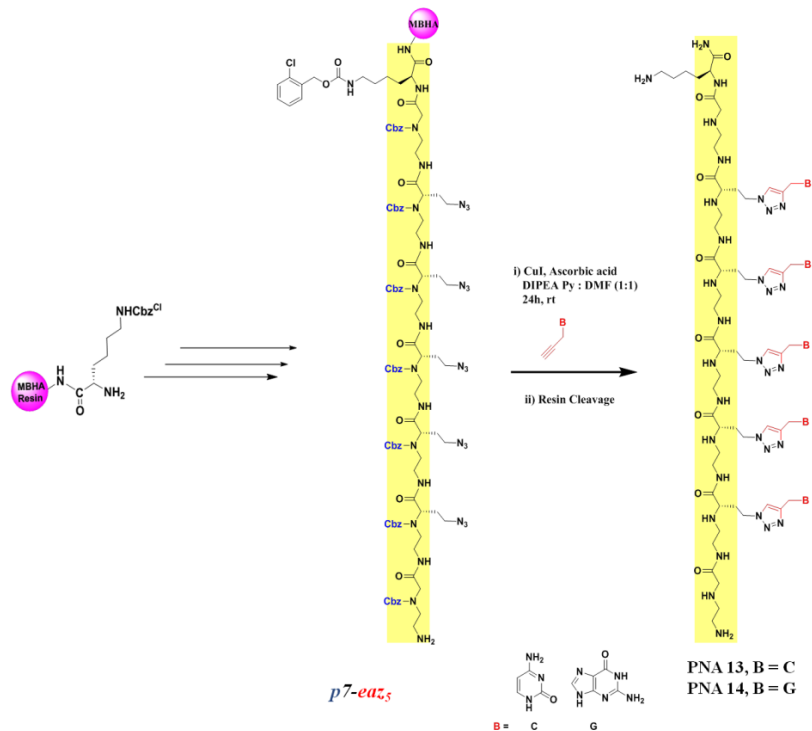
Figure 3.14 Strategy for solid phase synthesis of aminoethyl glycyl C$_\alpha$ triazolyl PNA (*aeg-tz* PNA) oligomers (PNA **13**, **PNA 14**)

***3.2.1f Summary of all PNA sequences*:** The sequences and structures of all PNA oligomers synthesized by solid phase in this chapter and the corresponding monomers used are summarized in Table 3.1.

Table 3.1 *Janus*-PNA oligomers with modified/unmodified monomers for base recognition

Entry	Sequence Code	PNA Sequences	Monomers used
1	T$_7$*jp-tz-C*$_5$ JP 1		α-(*S*-ethylazido) *aeg Thymine*
2	m*jp-tz-C*$_5$ JP 2		α-(*S*-ethylazido) *aeg Cytosine*
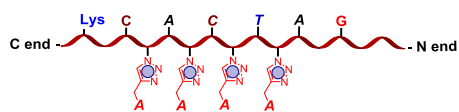
3	**m***jp-tz-**A**$_4$ **JP 3**		**α-(*S*-ethylazido)** *aeg Guanine*
4	**m***jp-tz-**G**$_4$ **JP 4**		**α-(*S*-ethylazido)** *aeg Adenine*
5	**m***jp-tz*$_4$**-**CACG**** **JP 5**		**α(*S*)-[(etz-N9(A^{NHCbz})]-N(acetamido thymnyl** *aeg*
6	**m***jp-tz*$_5$**-**CCACG**** **JP 6**		
7	***SCM*-**JP** JP 7**		
8	**T**$_7$***jp-**eaz**$_5$ **PNA 8**		
9	**m***jp-**eaz**$_5$ **PNA 9**		
10	**T**$_5$***jp-tz-**C**$_3$ **JP 10**		
11	C*A-G* **JP 11**		
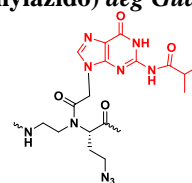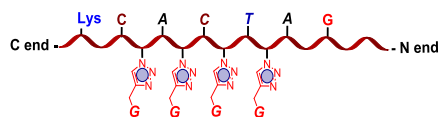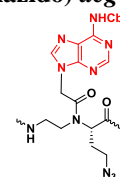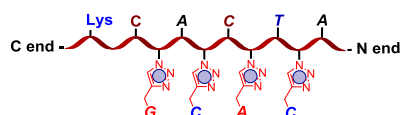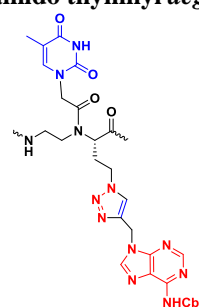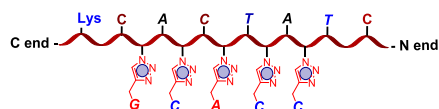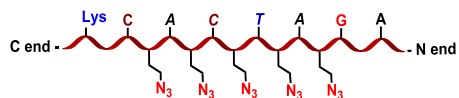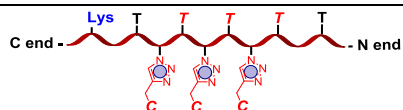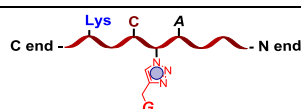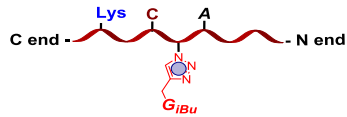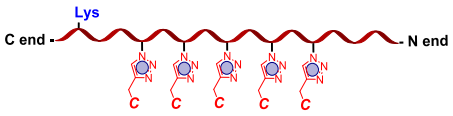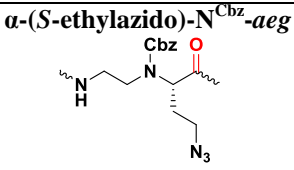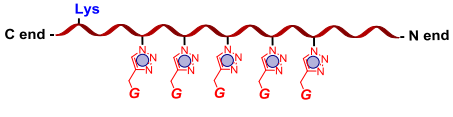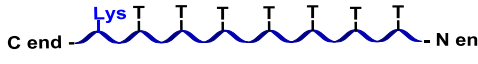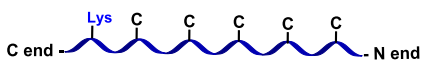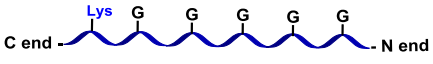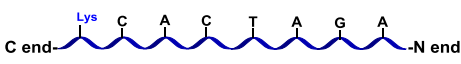
12	C*A*-*G*_{iBu} **JP 12**	
13	*p7*-*tz*-**C₅** **PNA 13**	
14	*p7*-*tz*-**G₅** **PNA 14**	
15	**PNA-T₇** **PNA 15**	
16	**PNA-C₅** **PNA 16**	
17	**PNA-G₅** **PNA 17**	
18	**PNA 18**	

3.2.2 Cleavage of PNA oligomers from the solid support

The synthesized *Janus* PNA oligomers were cleaved from the solid support (L-lysine derivatized MBHA resin) using trifluoromethane sulphonic acid (TFMSA) in the presence of trifluoroacetic acid (TFA). This yielded PNA oligomers having L-lysine amides at their C-termini.[11] In this cleavage condition, side chain protecting group was also removed accompanied by deprotection of nucleobases. After the cleavage reaction was over, the PNA oligomers obtained in solution were precipitated by addition of cold diethyl ether and the PNA oligomers were dissolved in de-ionized water.

3.2.3 Purification and characterization of the PNA oligomers

The completely deprotected crude PNA oligomer released into solution was checked by HPLC on reverse phase column. The crude HPLC showed several peaks, each one of it isolated and check by mass spectra. The desired product peak in the crude PNA oligomer identified by

mass spectroscopy was then purified by semi-preparative HPLC. The purified PNA is rechecked by analytical HPLC and characterized by high resolution mass spectral measurements. These were used for biophysical studies.

After the cleavage from the solid support, all PNA oligomers were purified by reverse phase high performance liquid chromatography (RP-HPLC) on semi-preparative C18 column using a gradient system of acetonitrile-water. The purity of PNA oligomers was checked by reinjecting the sample on the same C18 semi-preparative column. All HPLC chromatograms of crude and pure peptides are shown in Appendix II.

MALDI-TOF mass spectrometry was used to confirm the integrity of the synthesized PNA Oligomers. In literature, various matrices like sinapinic acid (3,5-dimethoxy-4-hydroxycinnamic acid), 2,5-dihydroxybenzoic acid (DHB), α-cyano-4-hydroxycinnamic acid (CHCA) etc. have been reported to record the MALDI-TOF spectra. Among these, DHB and CHCA were used here as a matrix to record MALDI-TOF spectra for all the synthesized PNAs. The calculated as well as observed molecular weights for all the PNAs with their molecular formulas, the MALDI-TOF data for confirmation of mixed purine-pyrimidine PNA oligomers are shown in Table 3.2.

Table 3.2 MALDI-TOF spectral analysis of the synthesized *Janus* PNA/*aeg-tz* PNA/*aeg*-PNA oligomers

Sr. No.	PNA sequence Code	Molecular Formula	Calculated Mass	Observed Mass	Retention Time (min)
1	T_7*jp-tz-C_5* JP 1	$C_{128}H_{163}N_{61}O_{34}$	3100.10 $[M]^+$	3100.60	14.07
2	m*jp-tz-C_5* JP 2	$C_{128}H_{129}N_{75}O_{26}$	3185.33 $[M + Na]^+$	3185.48	13.72
3	m*jp-tz-A_4* JP 3	$C_{112}H_{136}N_{70}O_{19}$	2766.79 $[M]^+$	2767.00	14.30
4	m*jp-tz-G_4* JP 4	$C_{112}H_{136}N_{70}O_{23}$	2830.78 $[M]^+$	2830.70	13.70
5	m*jp-tz$_4$-CACG* JP 5	$C_{99}H_{124}N_{59}O_{19}$	2444.47 $[M + H]^+$	2444.56	9.20
6	m*jp-tz$_5$-CCACG* JP 6	$C_{129}H_{160}N_{74}NaO_{27}$	3200.33 $[M + Na]^+$	3201.21	14.33
7	*SCM-JP* JP 7	$C_{129}H_{160}N_{75}O_{28}$	3209.33 $[M+H]^+$	3209.17	13.99
8	T_7*jp-eaz$_5$* PNA 8	$C_{93}H_{128}N_{46}KO_{29}$	2393.43 $[M + K]^+$	2392.43	----

9	m*jp-eaz$_5$* PNA **9**	C$_{91}$H$_{122}$N$_{60}$O$_{20}$	2376.37 [M]$^+$	2380.85	----
10	**T$_5$***jp-tz-***C$_3$** JP **10**	C$_{88}$H$_{116}$N$_{41}$O$_{24}$	2132.15 [M]$^+$	2132.10	12.02
11	**C***A*-**G** JP **11**	C$_{37}$H$_{51}$N$_{23}$KO$_7$	969.06 [M + K]$^+$	971.54	12.30
12	**C***A*-**G**$_{iBu}$ JP **12**	C$_{41}$H$_{57}$N$_{23}$O$_8$	1000.06 [M]$^+$	1000.76	15.26
13	*p7-tz-***C$_5$** PNA **13**	C$_{79}$H$_{121}$N$_{47}$O$_{13}$	1937.15 [M]$^+$	1937.49	13.30
14	*p7-tz-***G$_5$** PNA **14**	C$_{84}$H$_{122}$N$_{57}$O$_{13}$	2138.28 [M + H]$^+$	2138.20	13.55
15	**PNA-T$_7$** PNA **15**	C$_{83}$H$_{113}$N$_{31}$NaO$_{29}$	2031.99 [M + Na]$^+$	2031.75	13.92
16	**PNA-C$_5$** PNA **16**	C$_{56}$H$_{80}$N$_{28}$O$_{16}$	1401.43 [M]$^+$	1401.81	12.00
17	**PNA-G$_5$** PNA **17**	C$_{83}$H$_{113}$N$_{31}$NaO$_{29}$	2031.99 [M + Na]$^+$	2031.75	11.57
18	**Mix PNA** PNA **18**	C$_{83}$H$_{109}$N$_{45}$O$_{21}$	2071.88 [M]$^+$	2072.41	12.98

3.3 Summary

The rationally designed modified PNA and *Janus* PNA monomers have been incorporated into 7-mer of *homo*, *chimeric*, *hetero* and *self-complementary mix Janus* PNA sequences. Modified monomers (**4** to **10**) incorporated at solid phase by using HOBt, HBTU and DIPEA as coupling reagent by solid phase peptide synthesis protocol. All the *Janus* PNA and unmodified *aeg* PNA oligomers obtained by solid phase synthesis were cleaved from solid support using appropriate protocol. The *Janus* PNA, *aeg-tz* PNA and *aeg* PNA oligomers after cleavage were purified by RP-HPLC and characterized by MALDI-TOF spectrometry. Detailed experimental procedures (sections 3.5.1, 3.5.2 and 3.5.3) and spectral data of all intermediates are discussed in sections 3.5.5 and 3.5.6). The next chapter deals with the investigation of biophysical properties of PNA oligomers.

3.4 Experimental Methods

3.5.1 Synthesis of *Janus* PNA oligomers on solid support by 'click' reaction

The *Janus*-PNA oligomers were synthesized by solid phase click reaction,[12-14] in which the resin-bound azide PNA oligomers (10 mg, 0.20 mmol/g) in DMF:pyridine (1:1) were reacted with corresponding nucleobase alkyne (9.0 mg, 6 eq) in the presence of CuI (12 mg, 18 equiv),

ascorbic acid (3.0 mg, 5 eq), and DIPEA (15 µL, 24 eq). The reaction was done for 5 min in microwave at 65 $^{\circ}$C and 25 W and then 24 h at room temperature. Excess reagents were removed by filtration and the resin was washed with DMF, DCM, MeOH and saturated EDTA.

3.4.2 Cleavage of the *Janus* PNA oligomers from solid support

The MBHA resins (10 mg) after assembly of *Janus* PNA oligomers were stirred with thioanisole (20 µL) and 1,2-ethanedithiol (8 µL) in an ice bath for 10 min. TFA (200 µL) was added and cooled in an ice bath. TFMSA (16 µL) was added slowly with stirring and the reaction mixture was stirred for 1.5 to 2 h at room temperature. The resin was removed by filtration under reduced pressure and washed twice with TFA and the filtrate was evaporated on a rotary evaporator at ambient temperature. The filtrate was transferred to eppendorf tube and the peptide was precipitated with cold dry ether. The peptide was isolated by centrifugation and the precipitate was dissolved in 40% methylamine and kept for another 8 h at room temperature for isobutyryl group deprotection from modified guanine derivative peptides. Then, the peptides were concentrated and filtered again and purified by HPLC.

3.4.3 Purification of the *Janus* PNA oligomers by RP-HPLC

The purification of PNAs was carried out on Dionex ICS 3000 HPLC system with semi-preparative BEH130 C18 (10×250 mm) column using solvents water and acetonitrile with composition A: 0.1% TFA in $CH_3CN:H_2O$ (5:95) and B= 0.1% TFA in $CH_3CN:H_2O$ (1:1). The gradient for elution was 100% A to 100% B in 20 min, with flow rate of 2 mL/min. The HPLC elutions were monitored at 220 and 254 nm wavelength.

3.5 References

1. Schnolzer, M.; Alewood, P.; Jones, A.; Alewood, D.; Kent, S. B. H. *Int. J. Pept. Protein Res.* **1992**, *40*, 180-193.

2. Kaiser, E.; Colescott, R. L.; Bossinger, C. D.; Cook, P. I. *Anal. Biochem.* **1970**, *34*, 595–598.
   18. Sarin, V. K.; Kent, S. B. H.; Tam, J. P.; Merrifield, R. B. *Anal. Biochem.* **1981**, *117*, 147-157.

3. Menchise, V.; De Simone, G.; Tedeschi, T.; Corradini, R.; Sforza, S.; Marchelli, R.; Capasso, D.; Saviano, M.; Pedone, C. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 12021–12026.

4. Nielsen, P. E.; Haaima, G.; Lohse, A.; Buchardt, O. *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 1939-1942.

5. Bose, T.; Banerjee, A.; Nahar, S.; Maiti, S.; Kumar, V. A. *Chem. Commun.* **2015**, *51*, 7693-7696.

6. Christensen, L.; Fitzpatrick, R.; Gildea, B.; Petersen, K. H.; Hansen, H. F.; Koch, T.; Egholm, M.; Buchardt, O. Nielsen, P. E.; Coull, J.; Berg, R. H. *J. Pept. Sci.* **1995**, *1*, 175–183.

7. (a) Kramer, R. A.; Bleicher, K. H.; Wennemers, H. *Helv. Chim. Acta* **2012**, *95*, 2621-2634.
   (b) Porcheddu, A.; Giacomelli, G.; Piredda, I.; Carta, M.; Nieddu, G. *Eur. J. Org. Chem.* **2008**, 5786-5797.

8. 4. Halay, E.; Ay, E.; Salva, E.; Ay, K.; Karayildirim, T. *Nucleos. Nucleot. Nucl.* **2017**, *36*, 598-619.

9. Nagapradeep, N.; Verma, Sandeep. *Chem. Commun.* **2011**, *47*, 1755-1757.

10. Christensen, L.; Fitzpatrick, R.; Gildea, B.; Petersen, K. H.; Hansen, H. F.; Koch, T.; Egholm, M.; Buchardt, O.; Nielsen, P. E.; Coull, J.; Berg, R. H. *J. Pept. Sci.* **1995**, *3*, 175–183.

11. Gasser, G.; Koster, N. S. D.; Metzler-Nolte, N. *Chem. Commun.* **2008**, 3675-3677.

12. Kuijpers, B. H. M.; Groothuys, S.; Hawner, C.; Dam, J. t.; Quaedflieg, P. J. L. M.; Schoemaker, H. E.; Delft, F. L. v.; Rutjes, F. P. J. T. *Org. Process Res. Dev.* **2008**, *12*, 503-511.

13. Tornoe, C. W.; Christensen, C.; Meldal, M. *J. Org. Chem.* **2002**, *67*, 3057-3064.

14. Chow, H. Y.; Zhang, Y.; Matheson, E.; Li, X. *Chem. Rev.* **2019**.

3.6 Appendix II

3.6.1 HPLC chromatograms of crude and purified peptides

HPLC chromatogram of (JP 1)
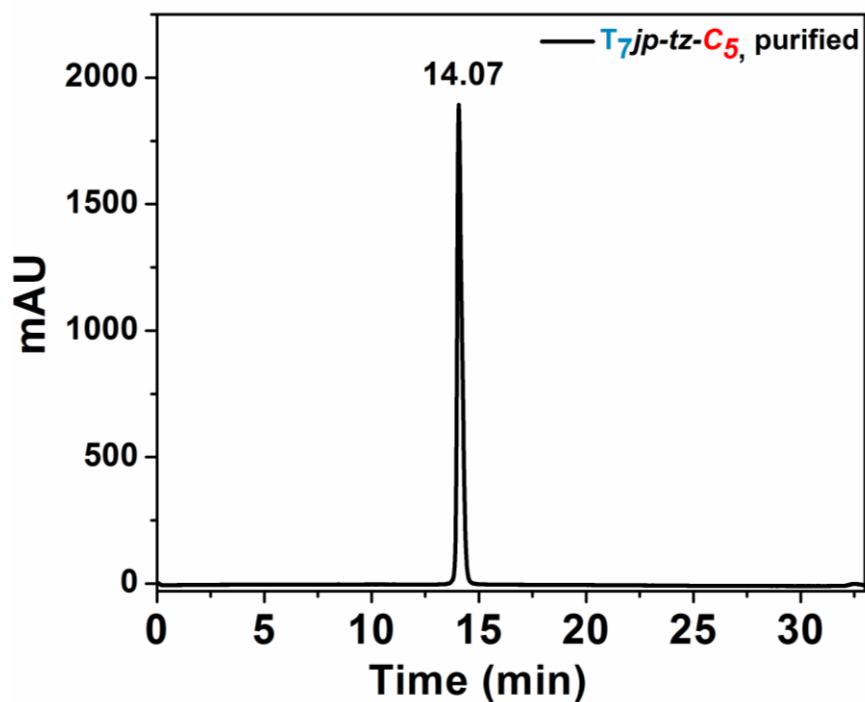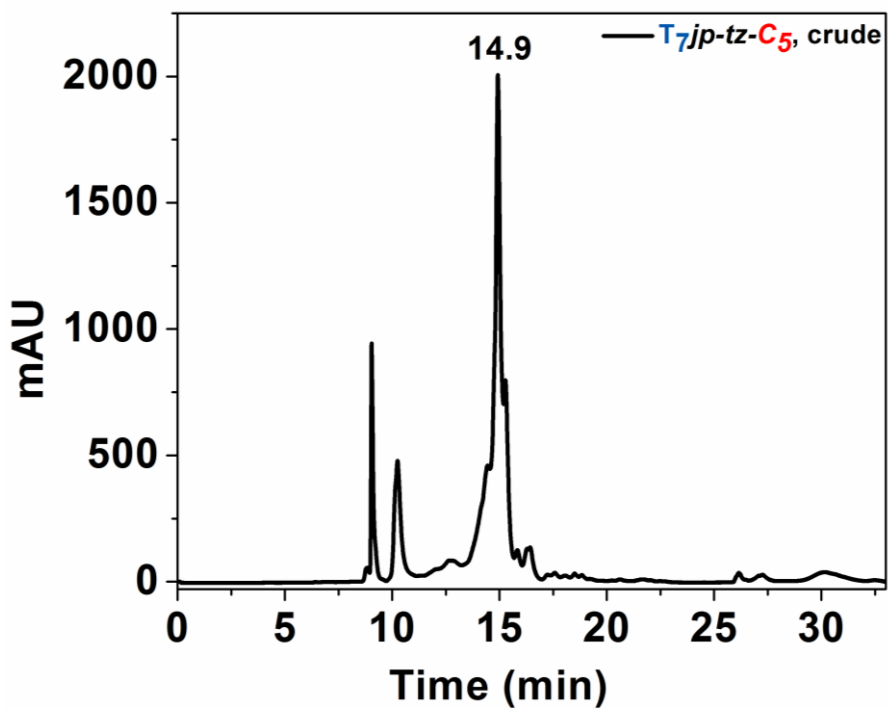
HPLC chromatogram of (JP 2)
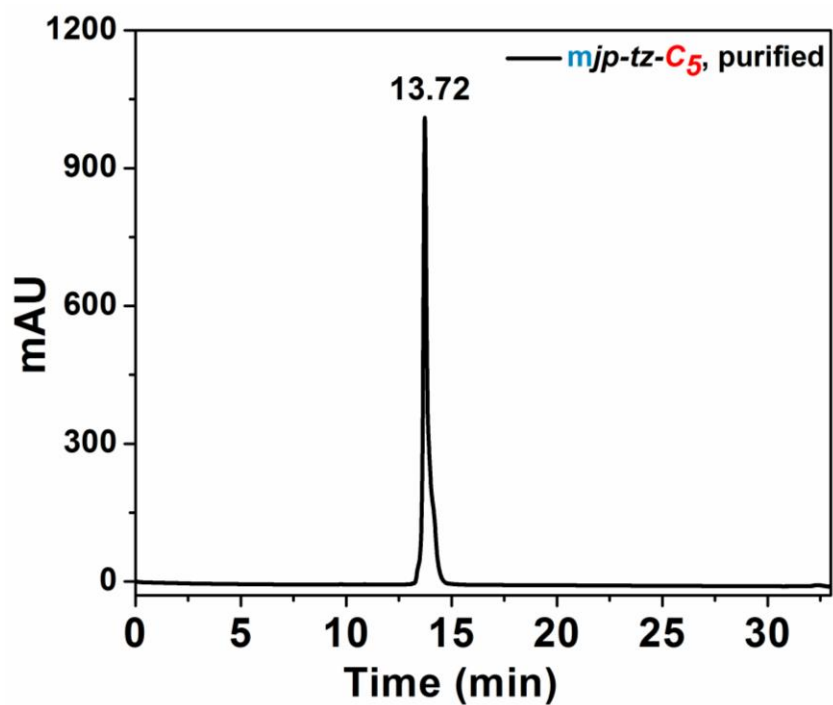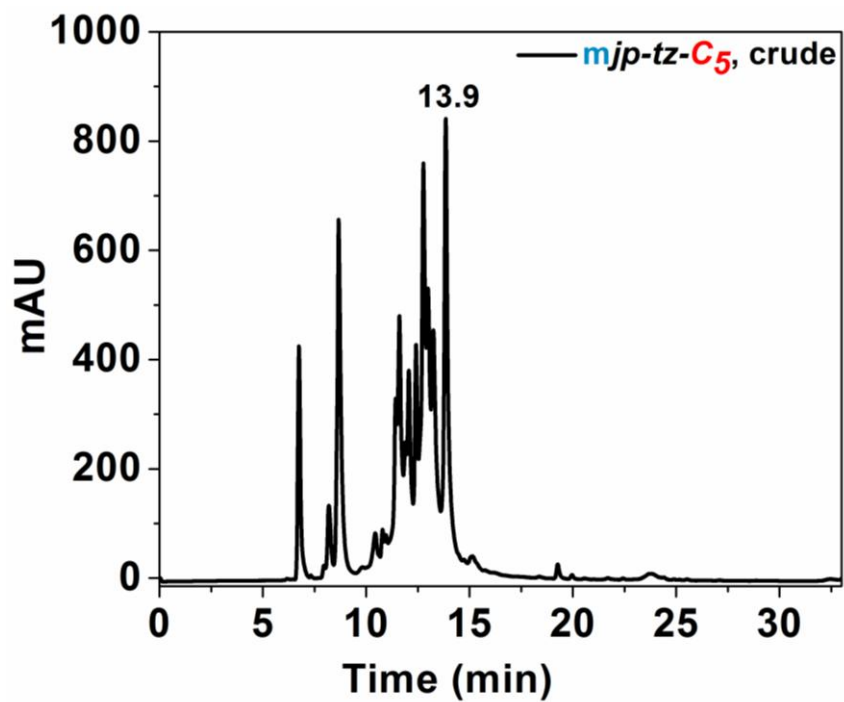
HPLC chromatogram of (JP 3)
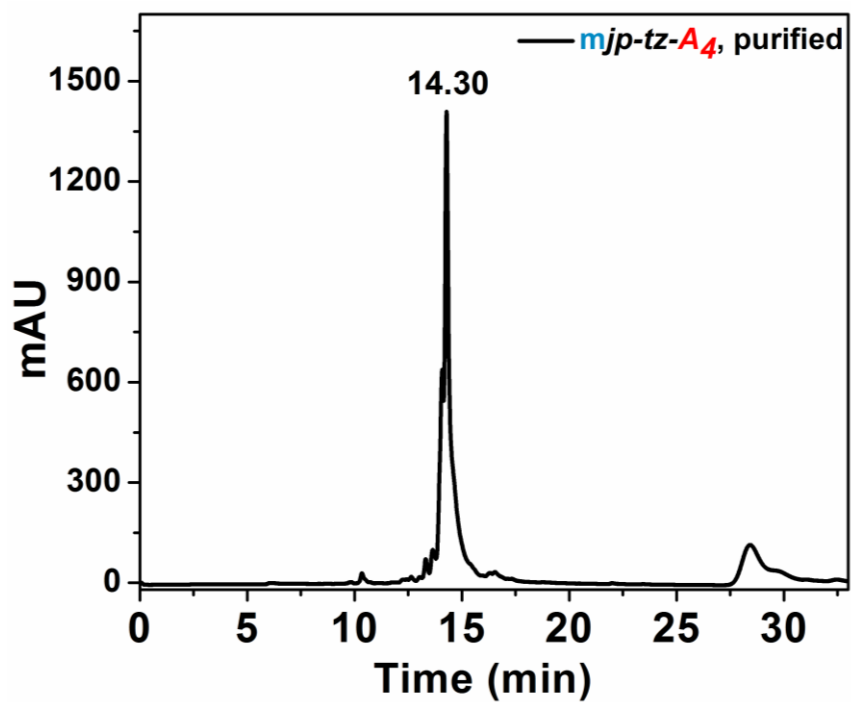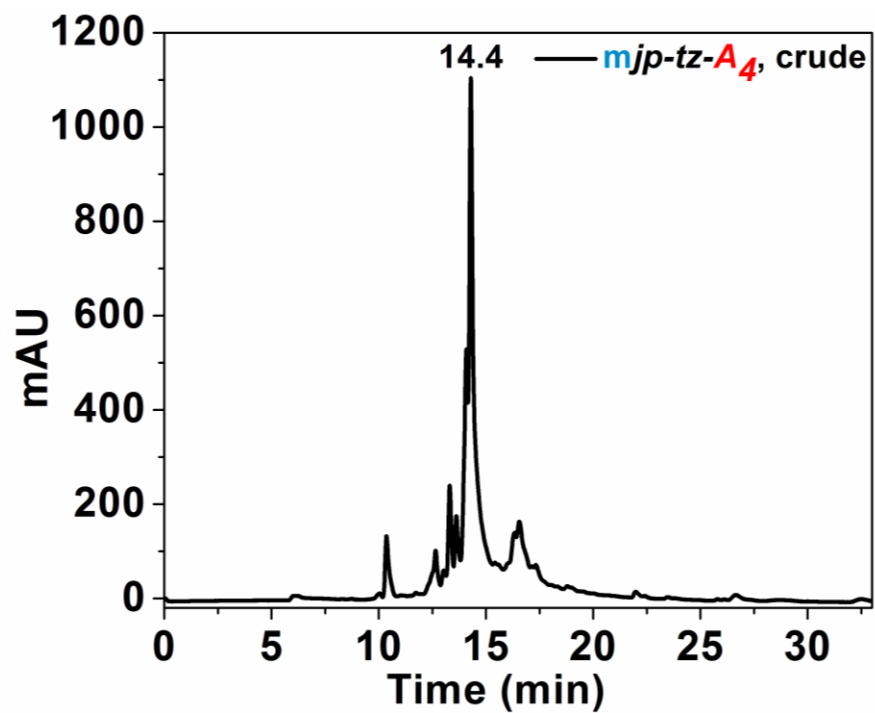
HPLC chromatogram of (JP 4)
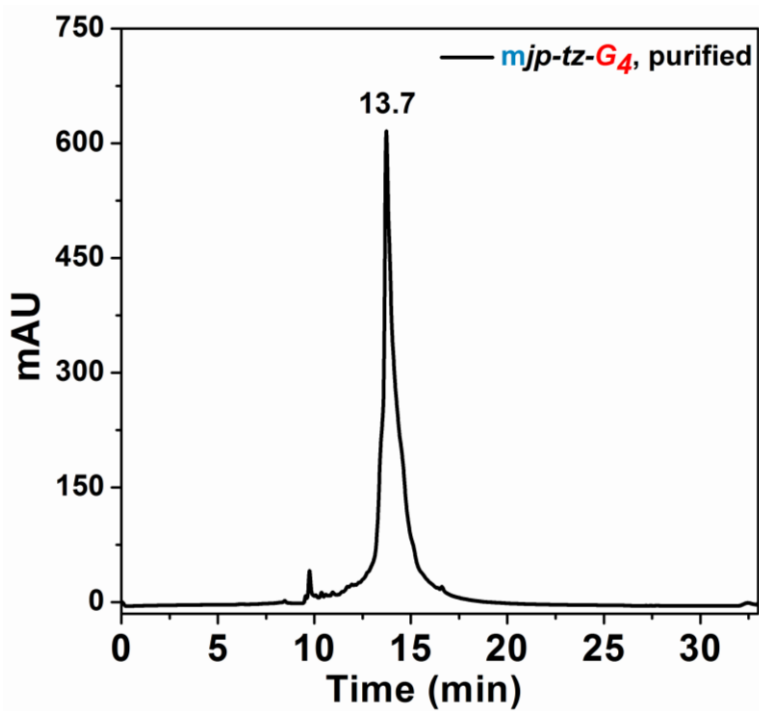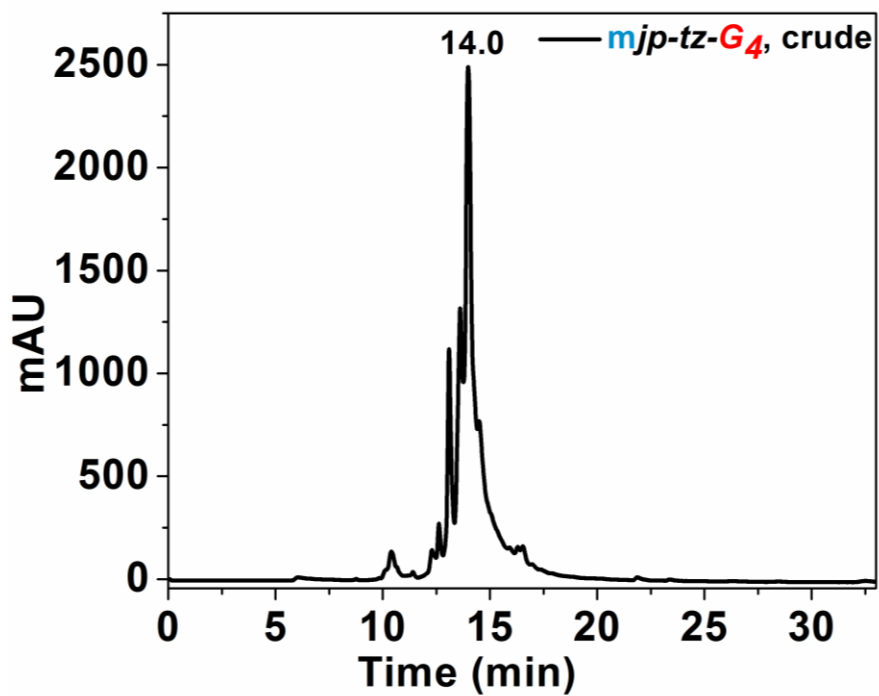
HPLC chromatogram of (JP 5)
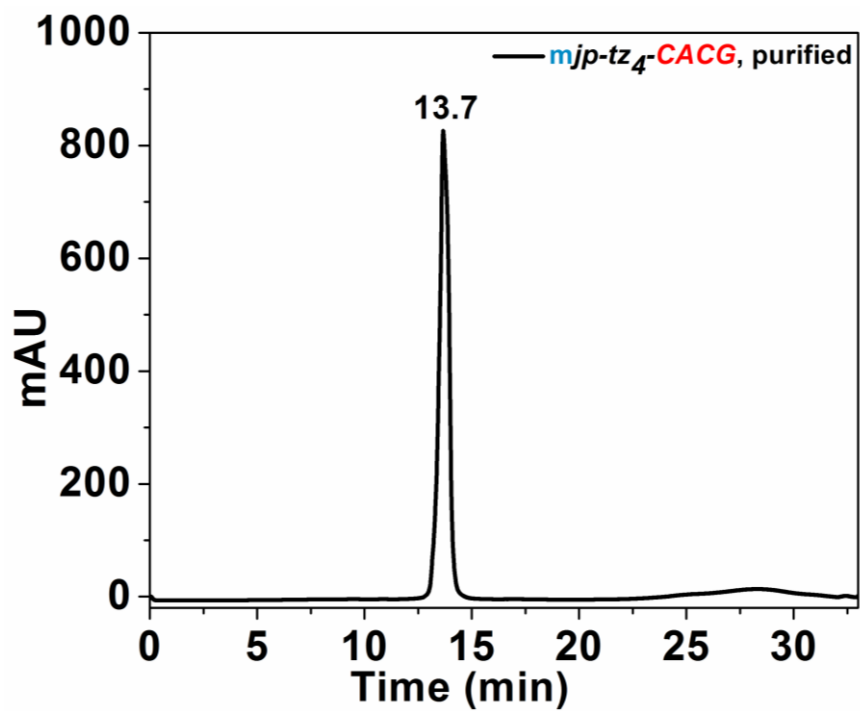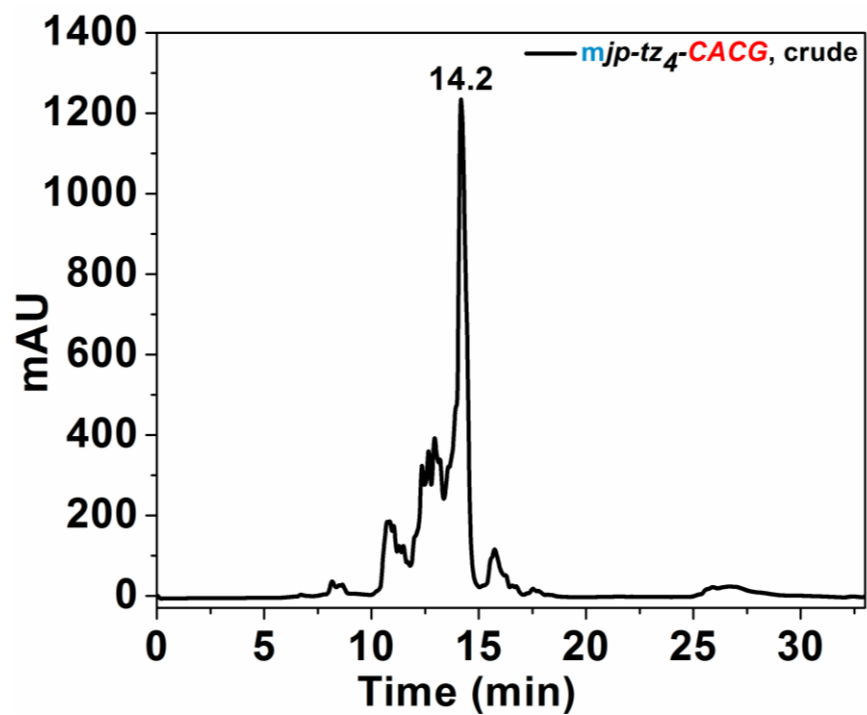
HPLC chromatogram of (JP 6)
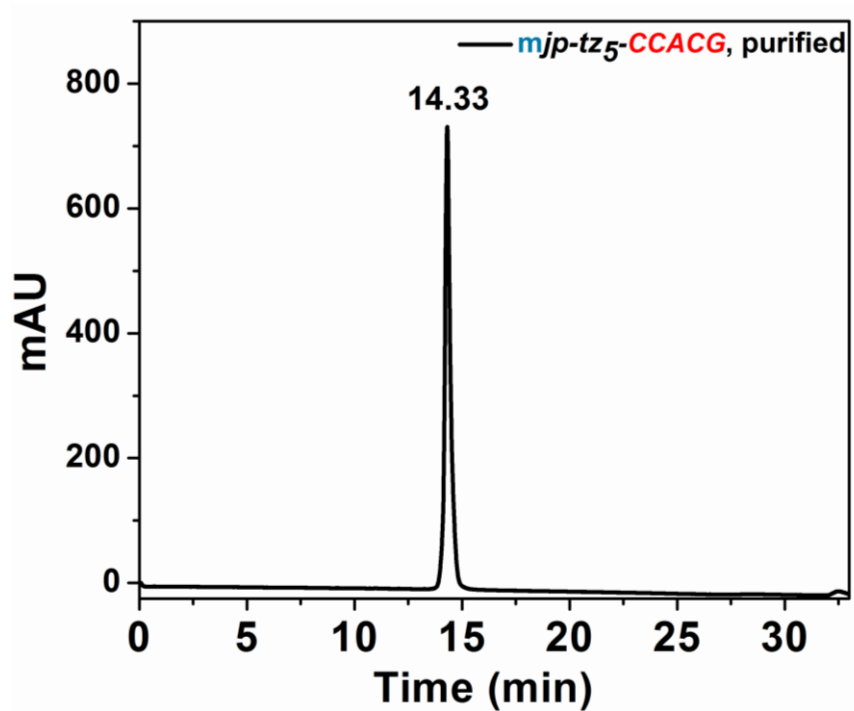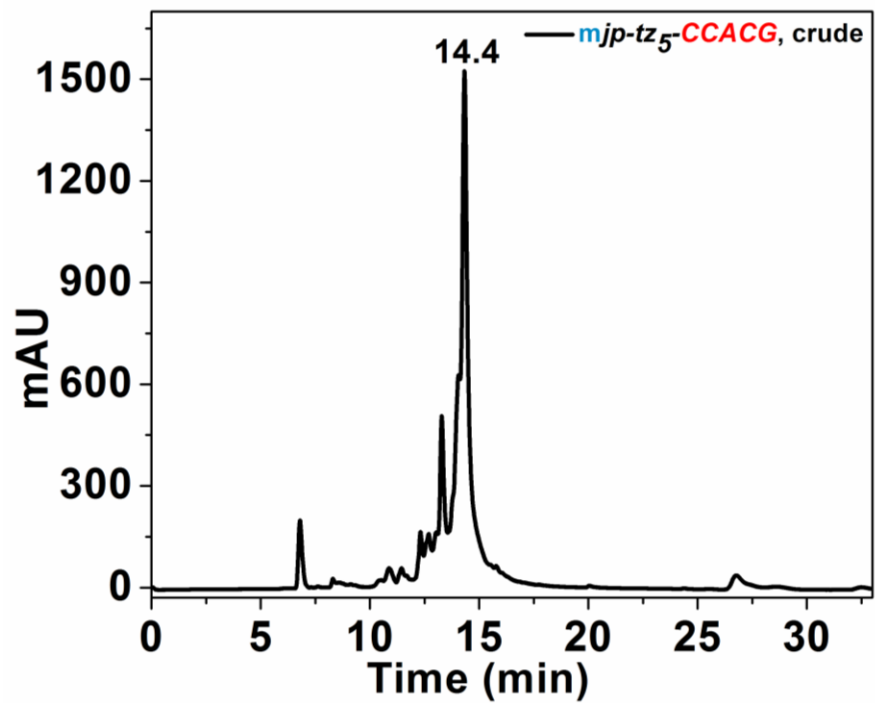
HPLC chromatogram of (JP 7)
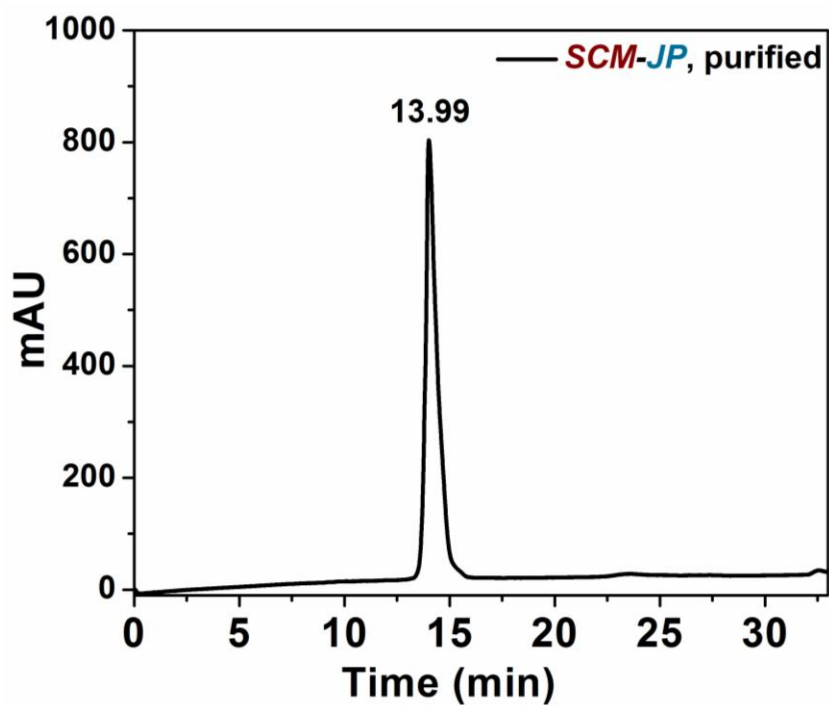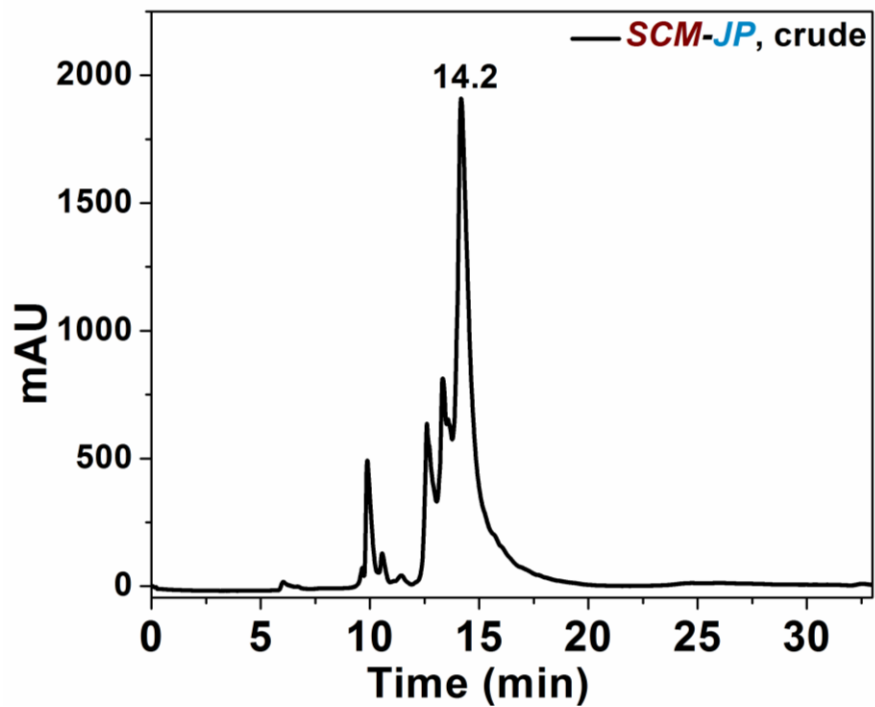
HPLC chromatogram of (JP 10)
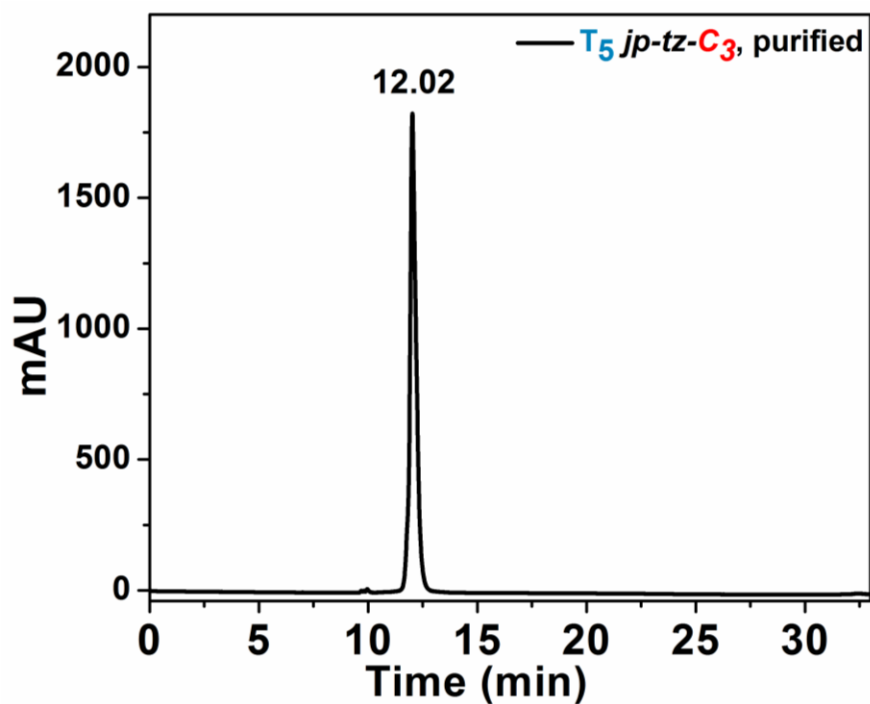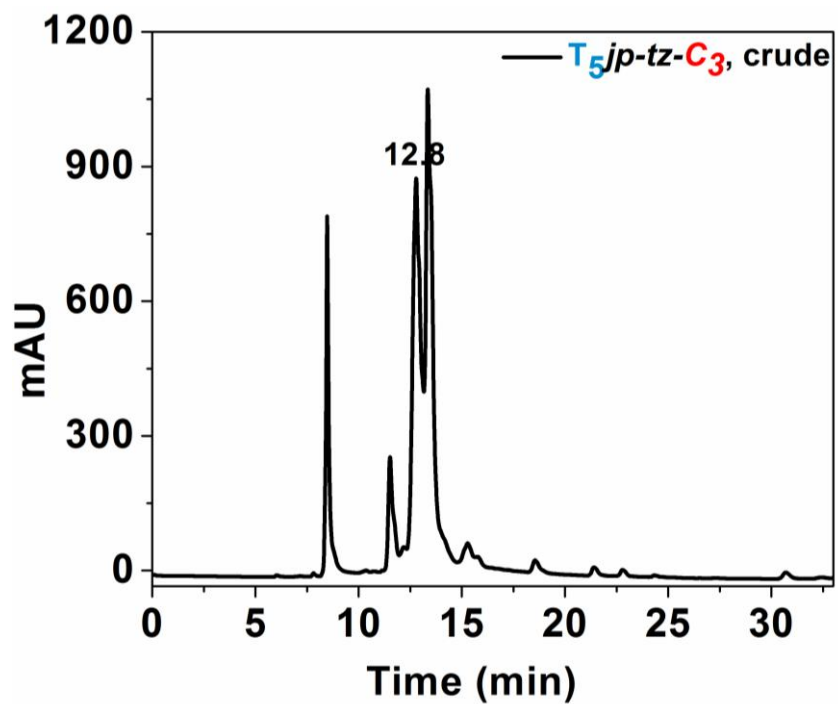
HPLC chromatogram of (JP 11)
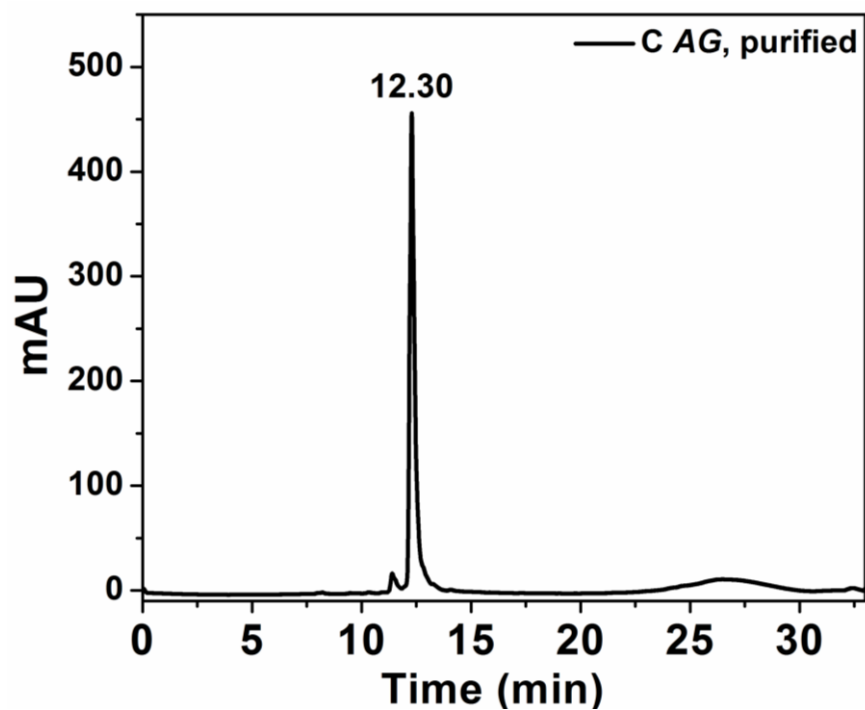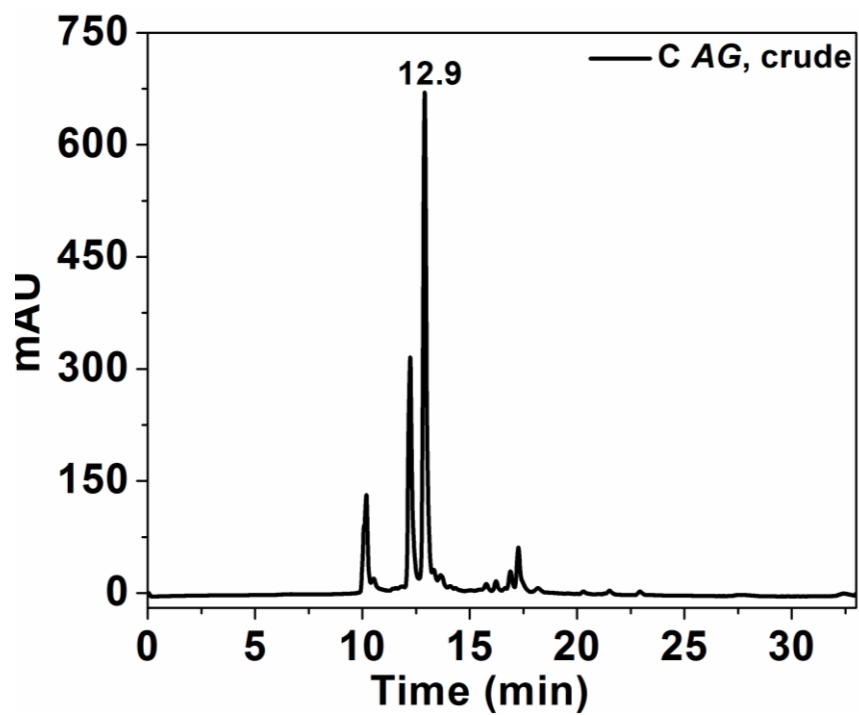
HPLC chromatogram of (JP 12)
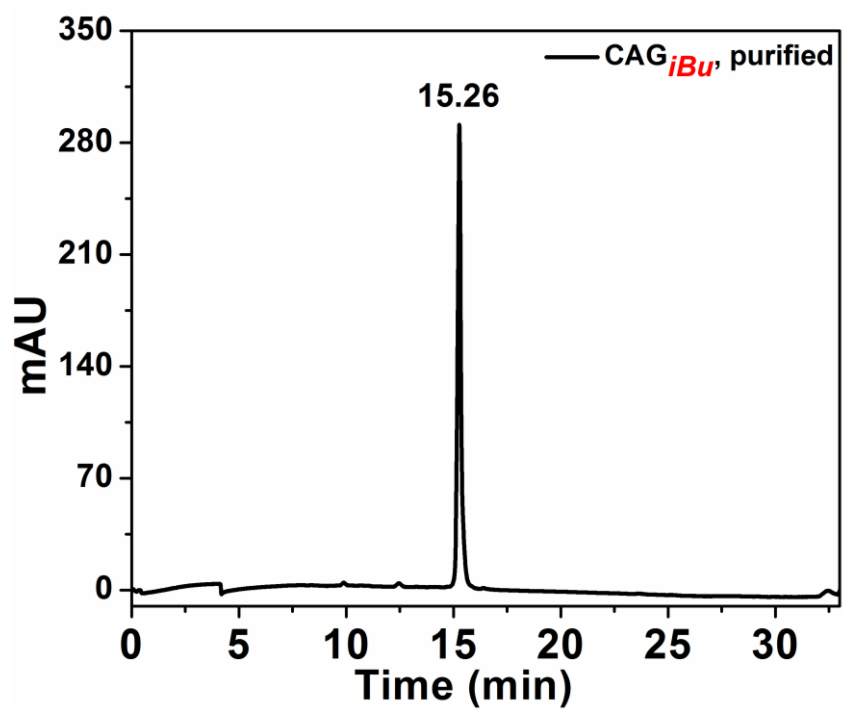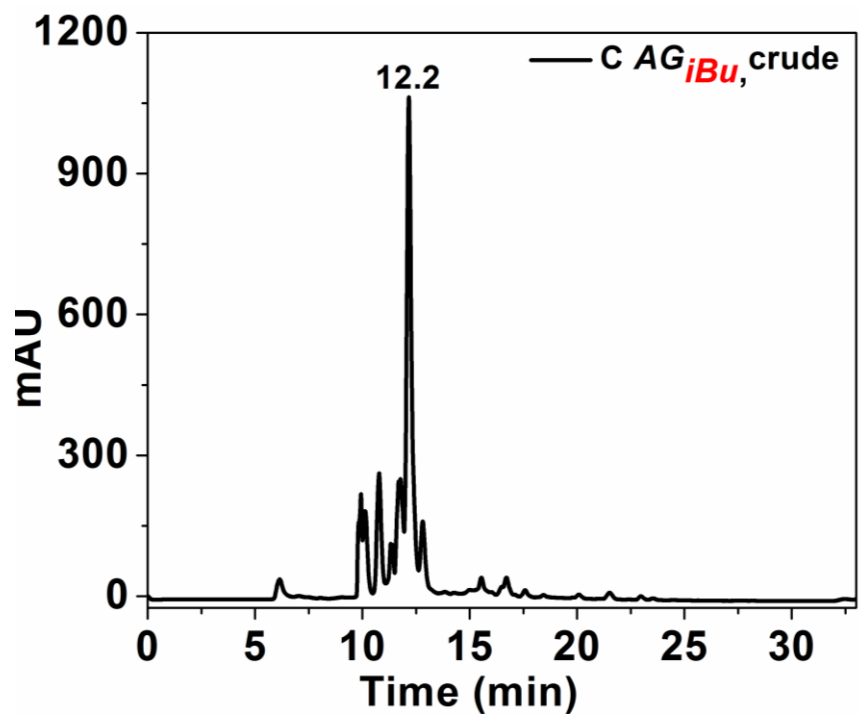
HPLC chromatogram of (PNA 13)
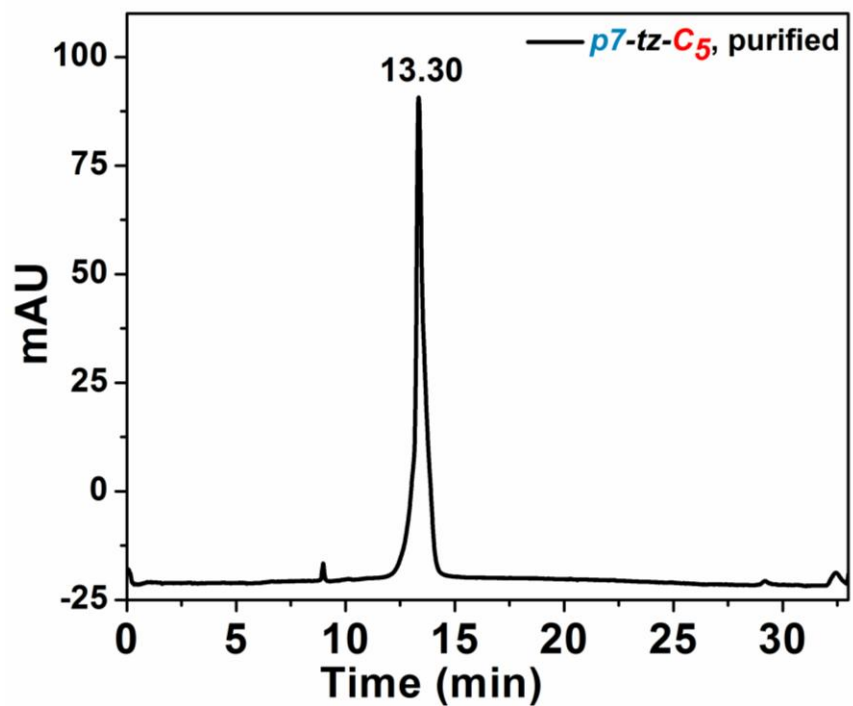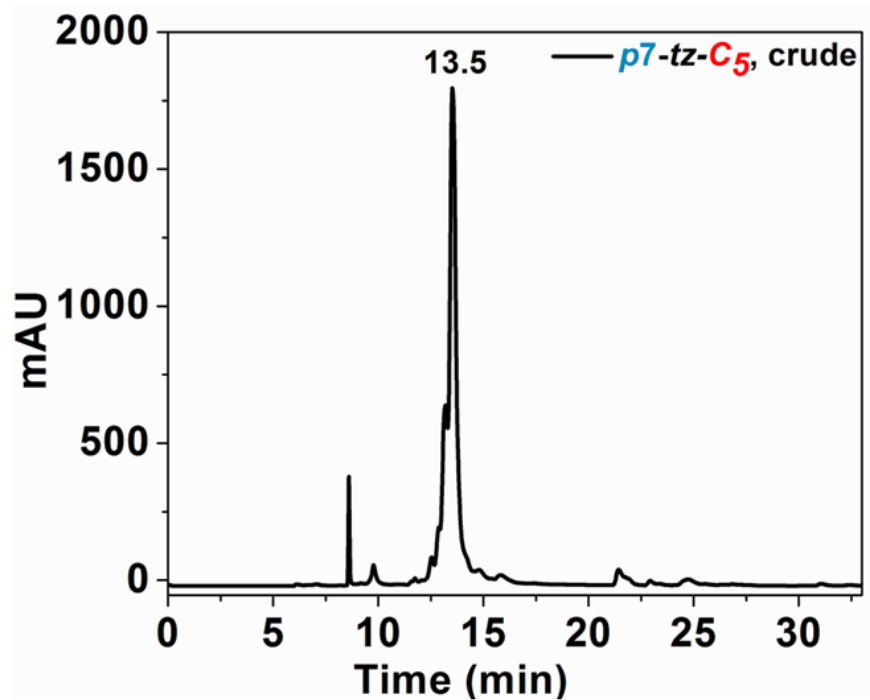
HPLC chromatogram of (PNA 14)
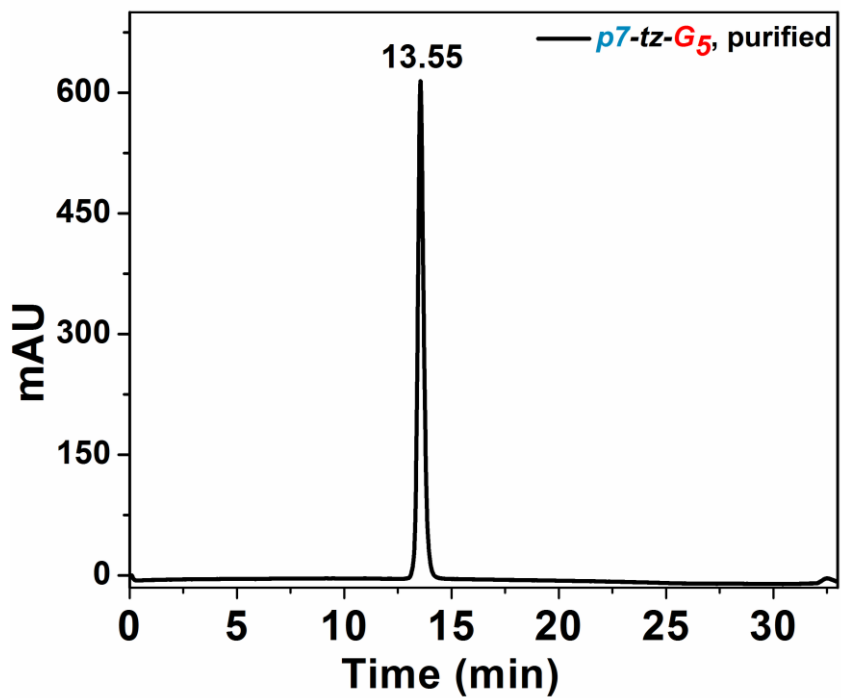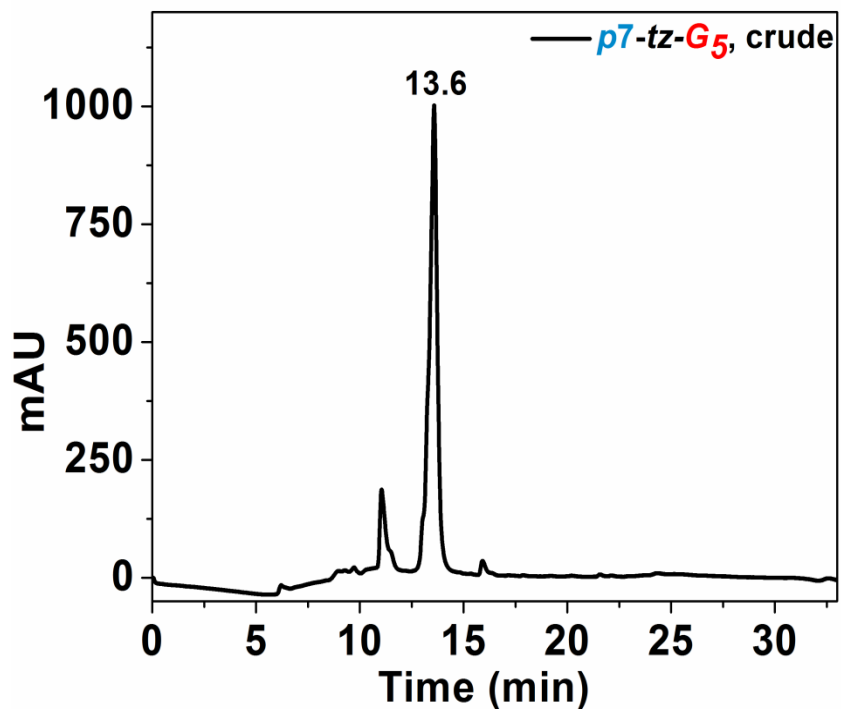
HPLC chromatogram of (PNA 15)
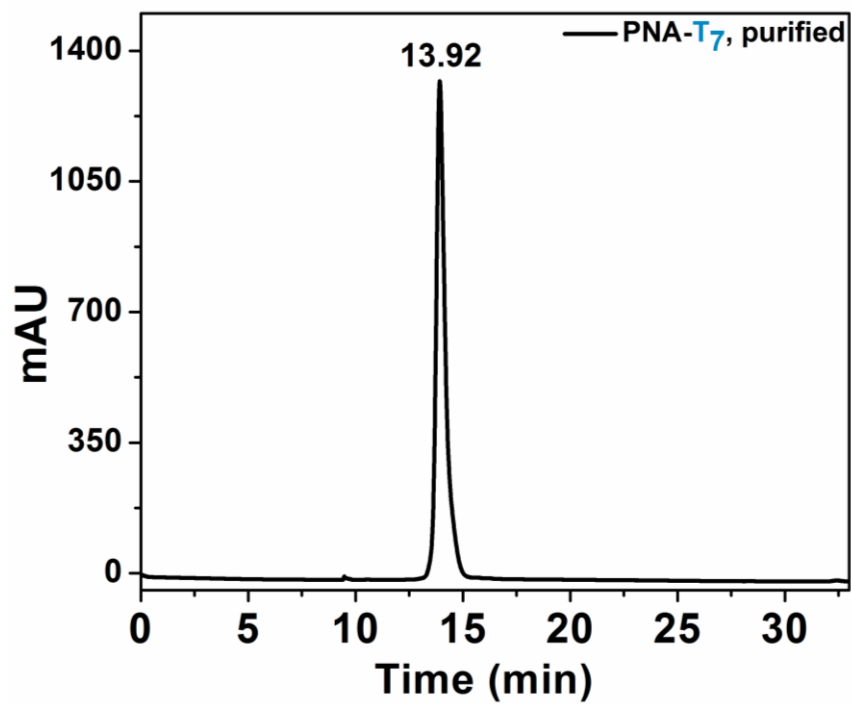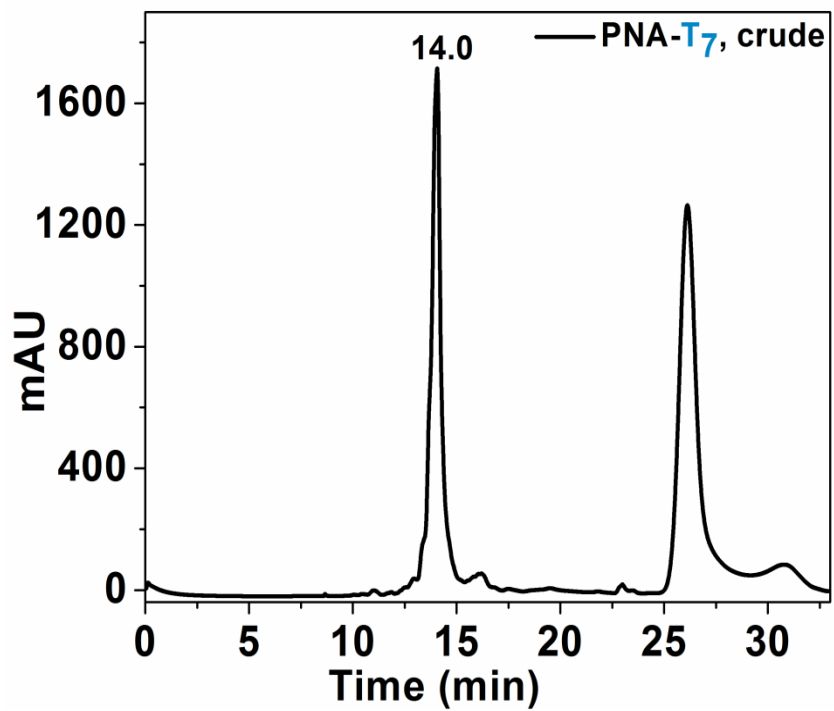
HPLC chromatogram of (PNA 16)
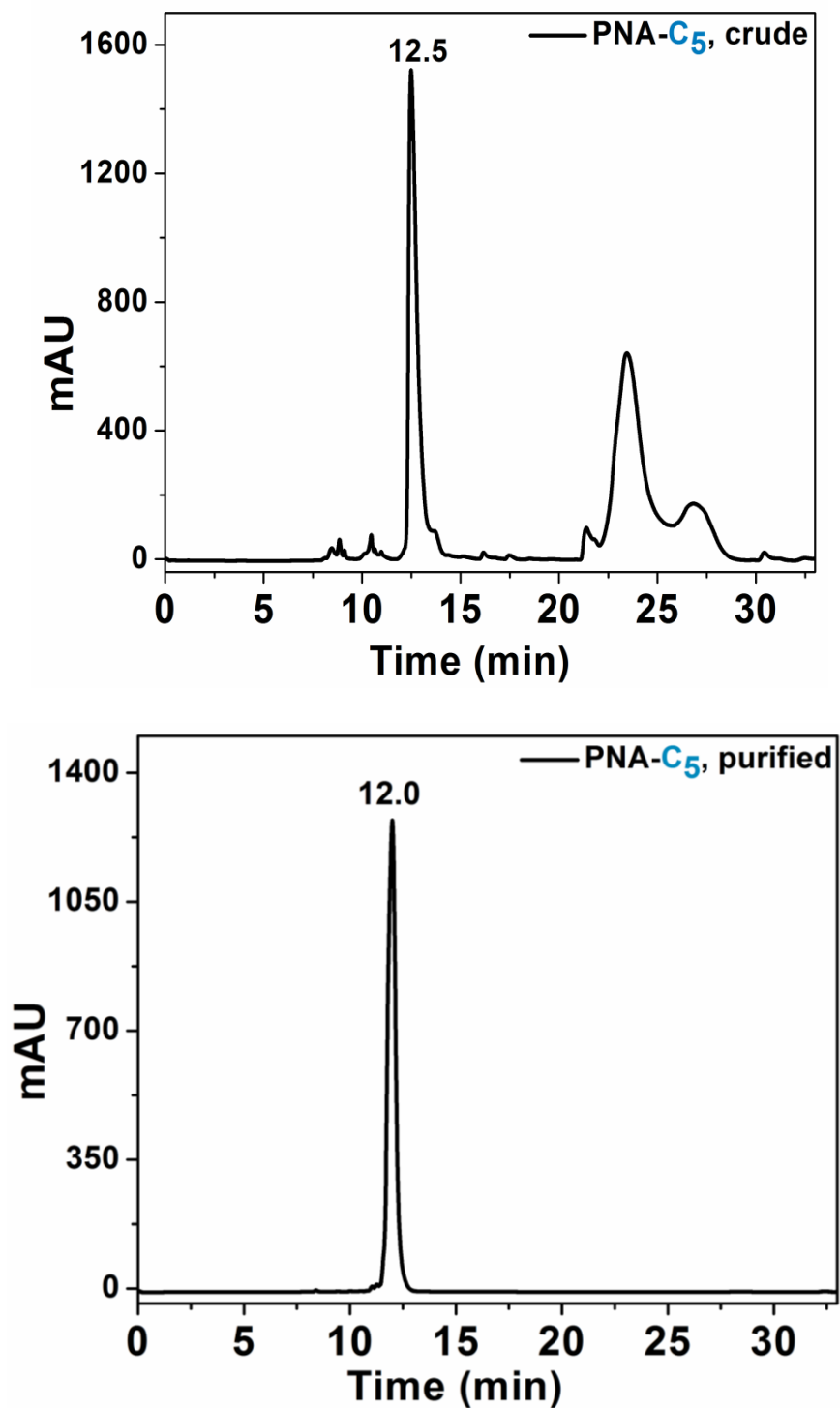
HPLC chromatogram of (PNA 17)
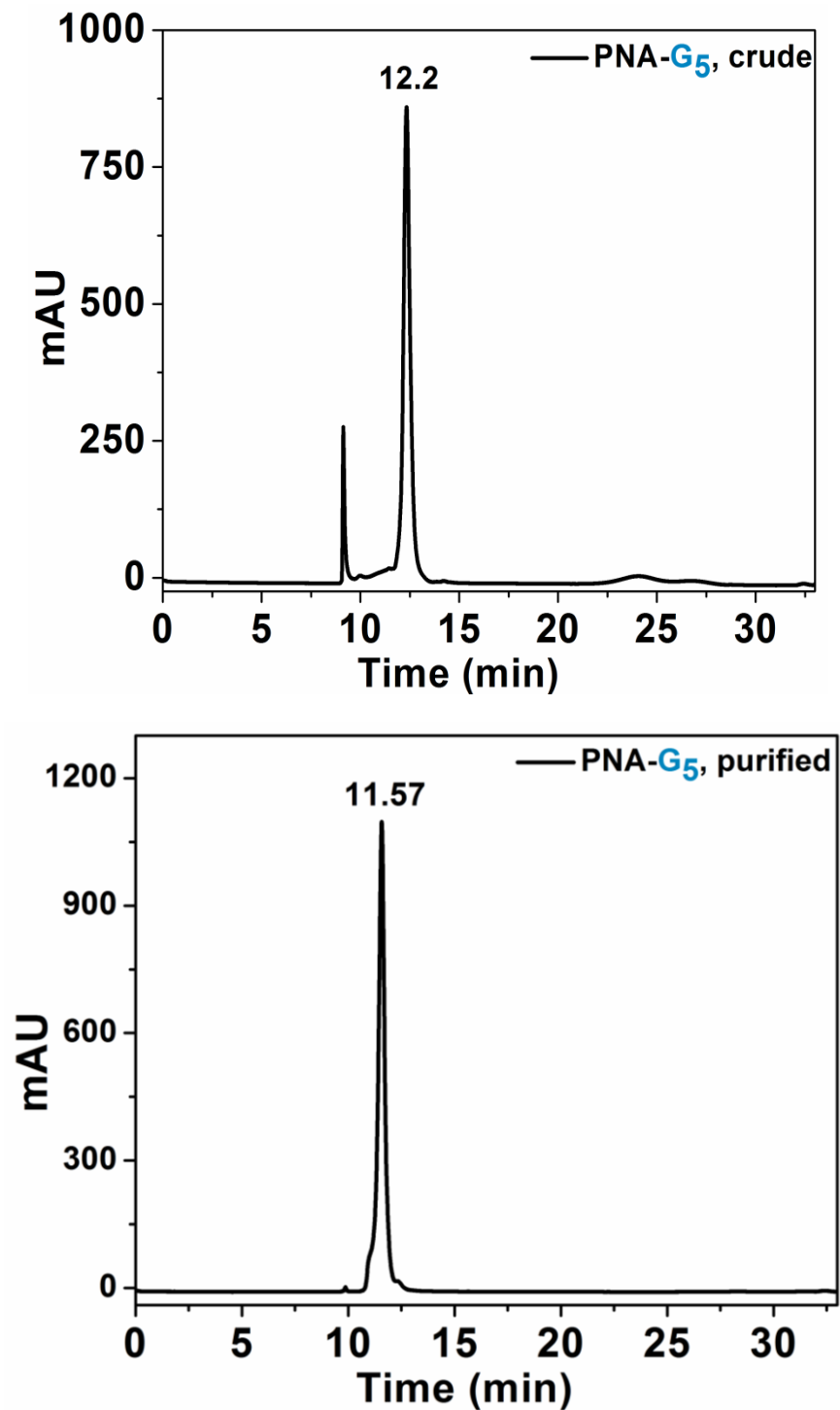
HPLC chromatogram of (PNA 18)
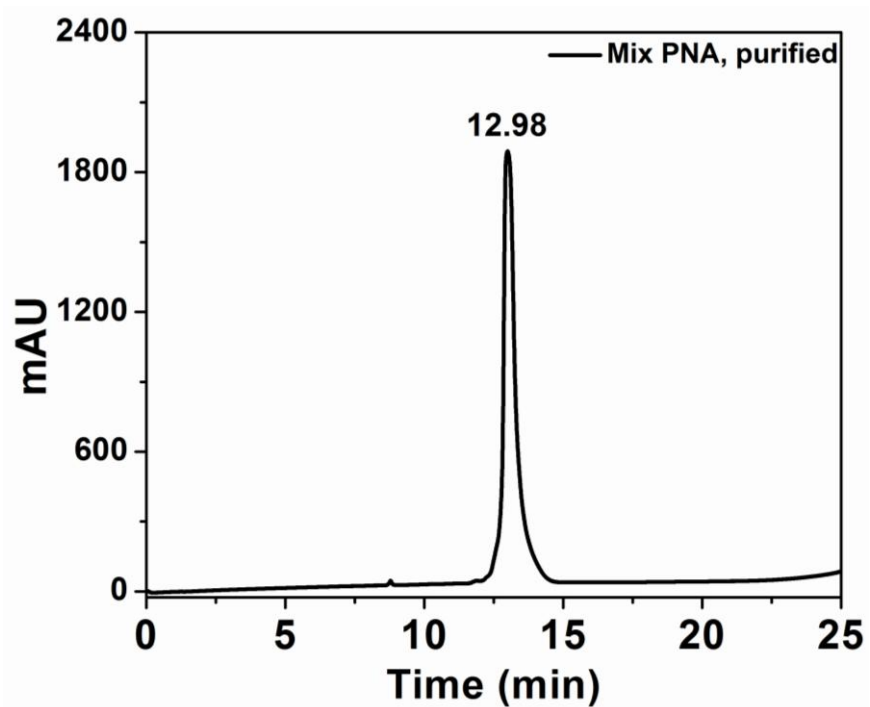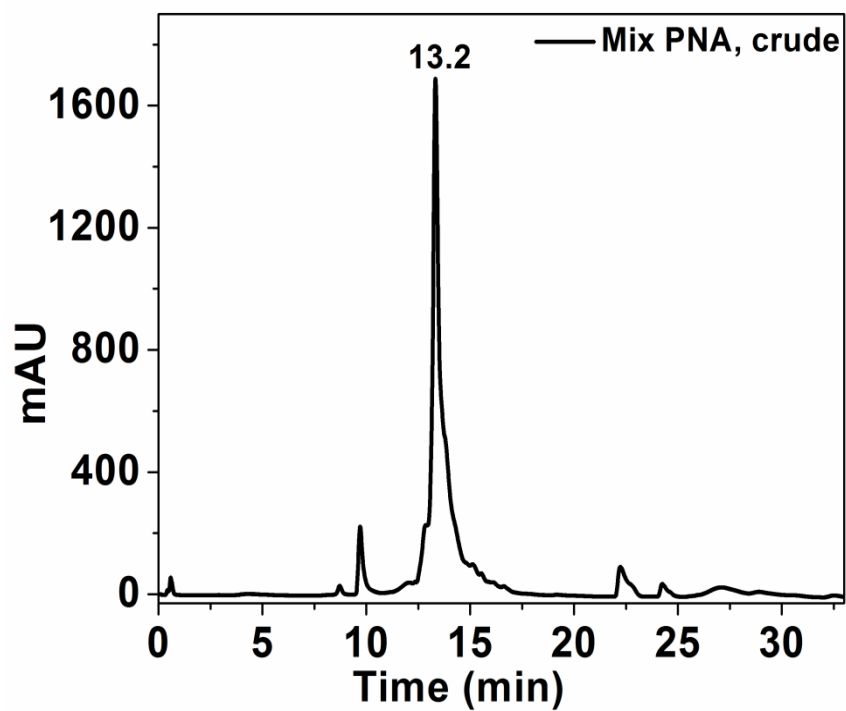
3.6.2 MALDI TOF spectra of peptides

MALDI-TOF Spectrum of (JP 1)
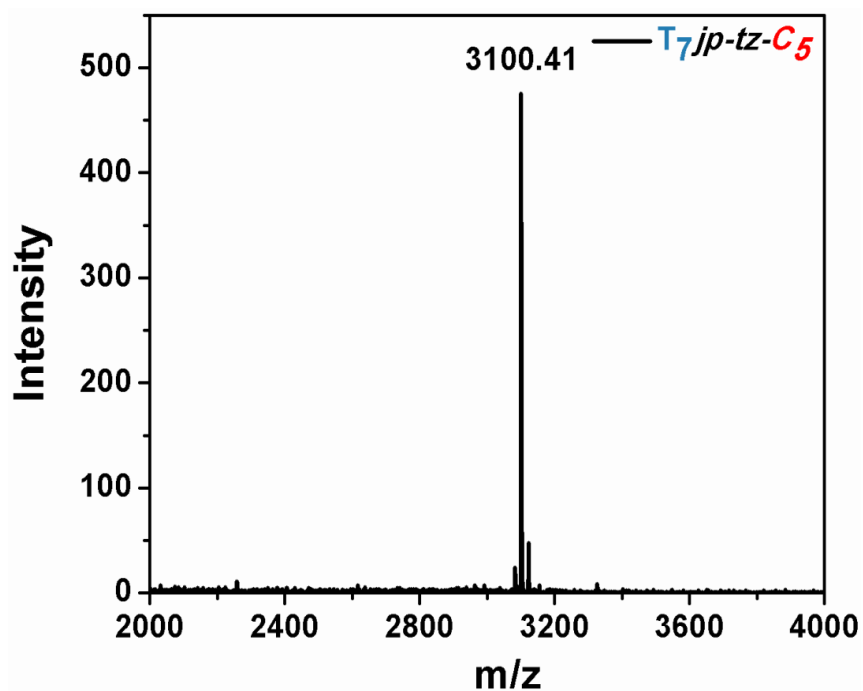
MALDI-TOF Spectrum of (JP 2)
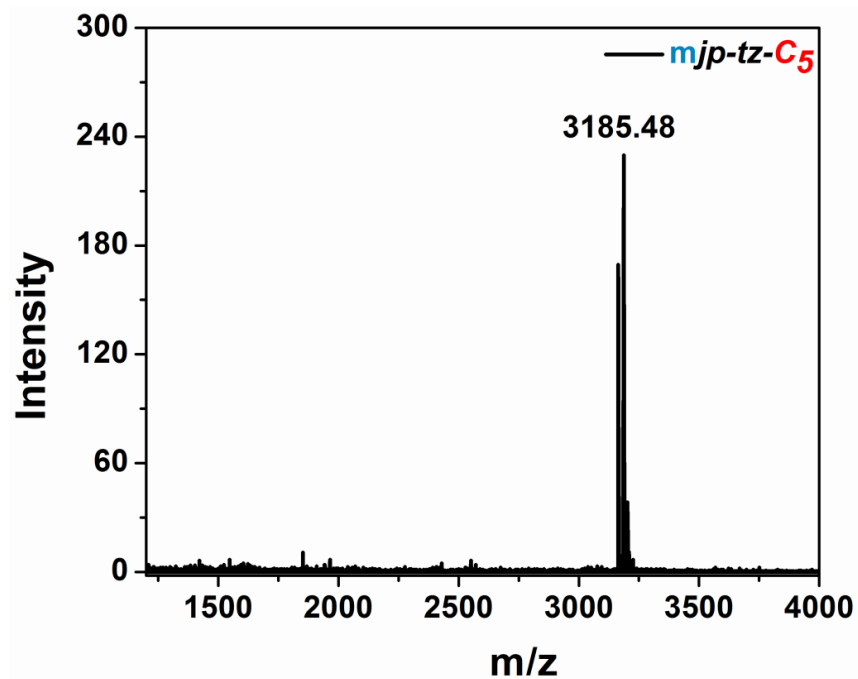
MALDI-TOF Spectrum of (JP 3)
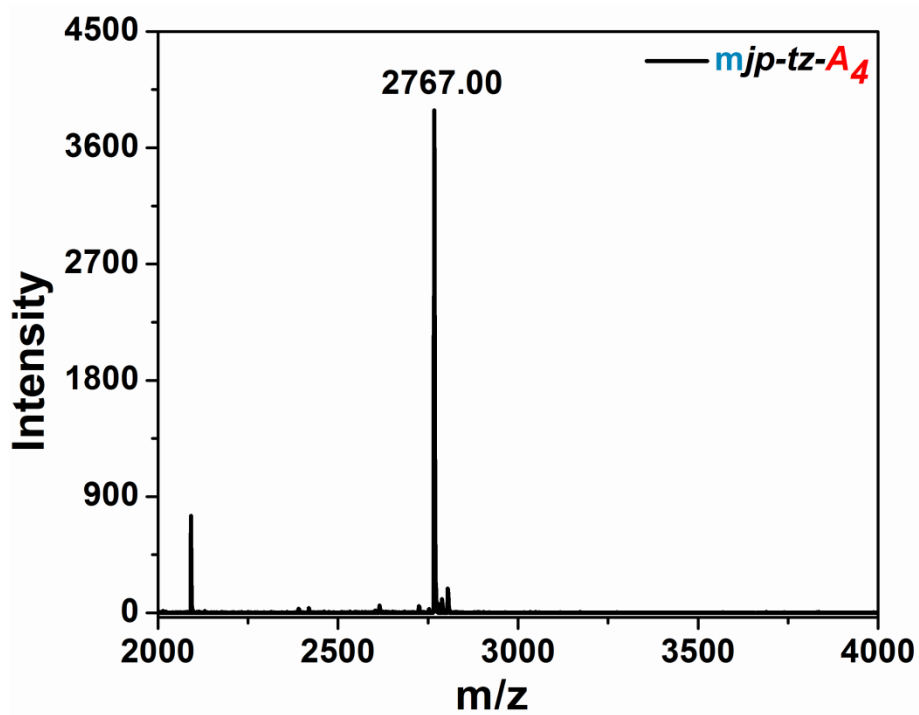
MALDI-TOF Spectrum of (JP 4)
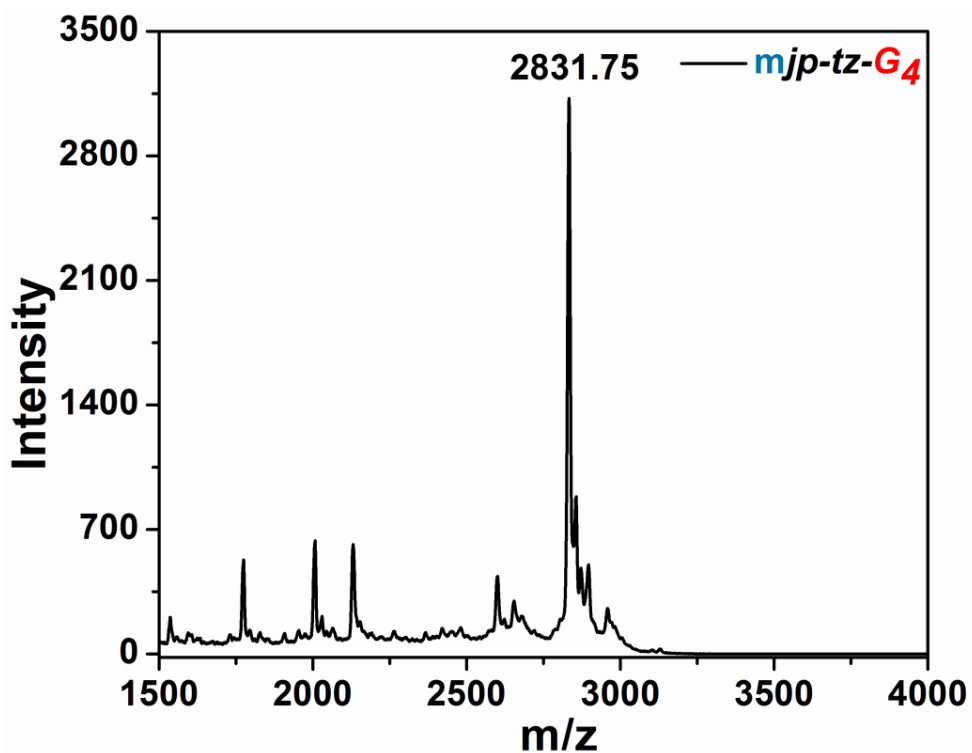
MALDI-TOF Spectrum of (JP 5)
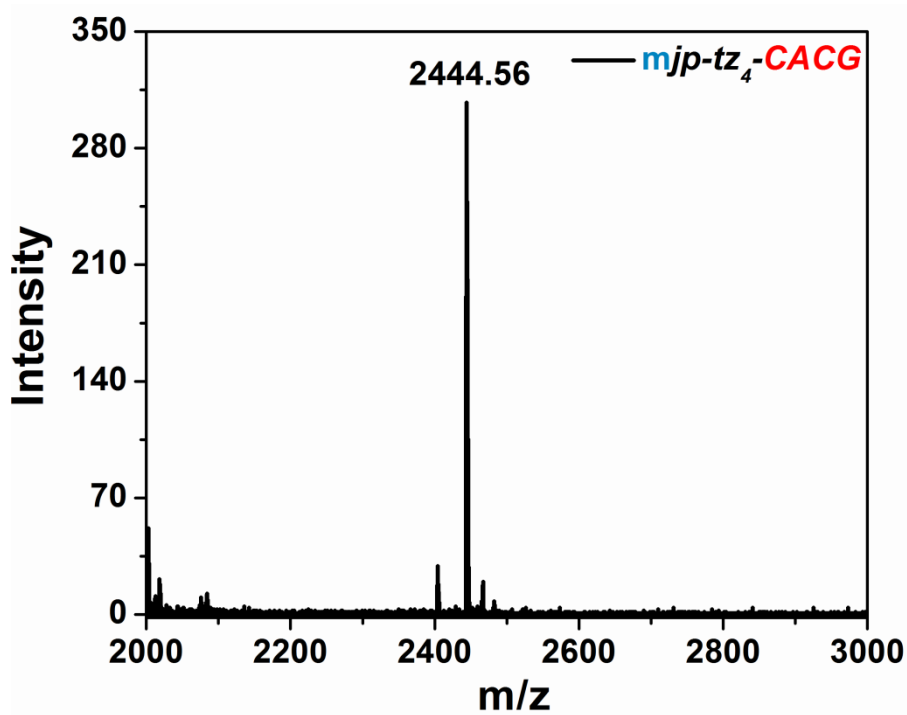
MALDI-TOF Spectrum of (JP 6)
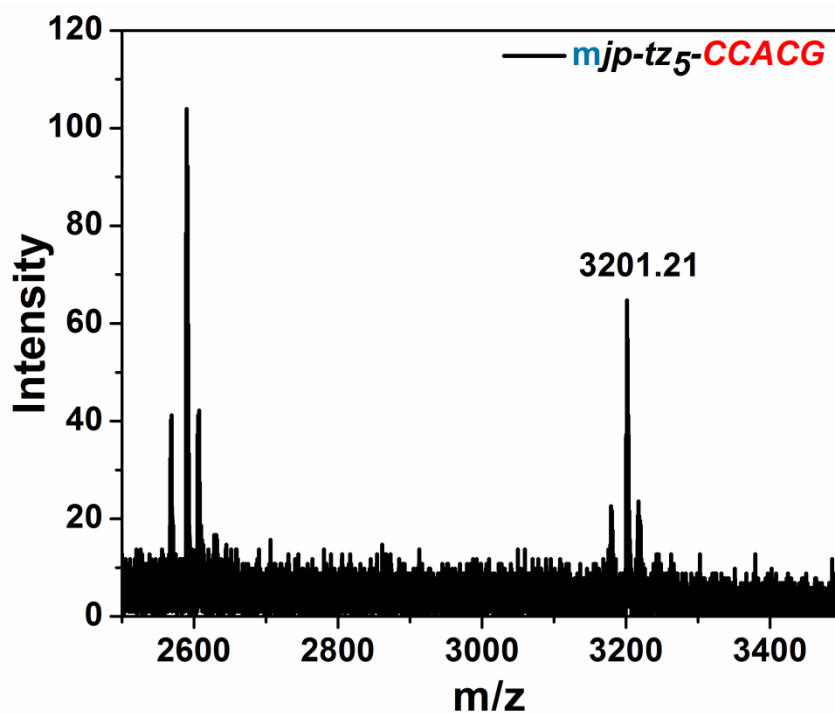
MALDI-TOF Spectrum of (JP 7)
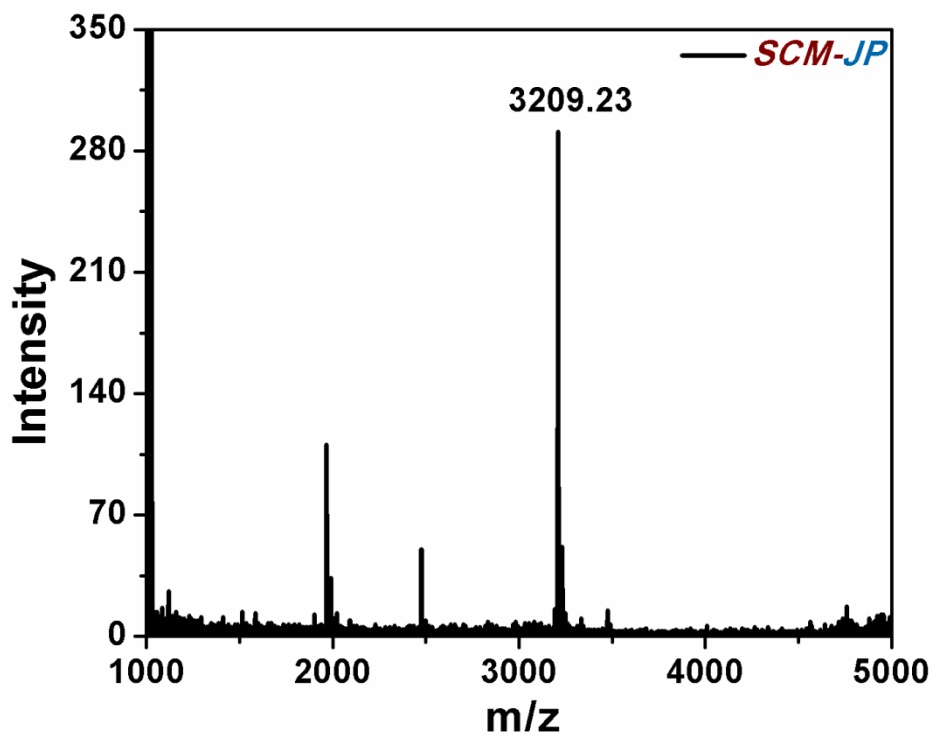
MALDI-TOF Spectrum of (PNA 8)
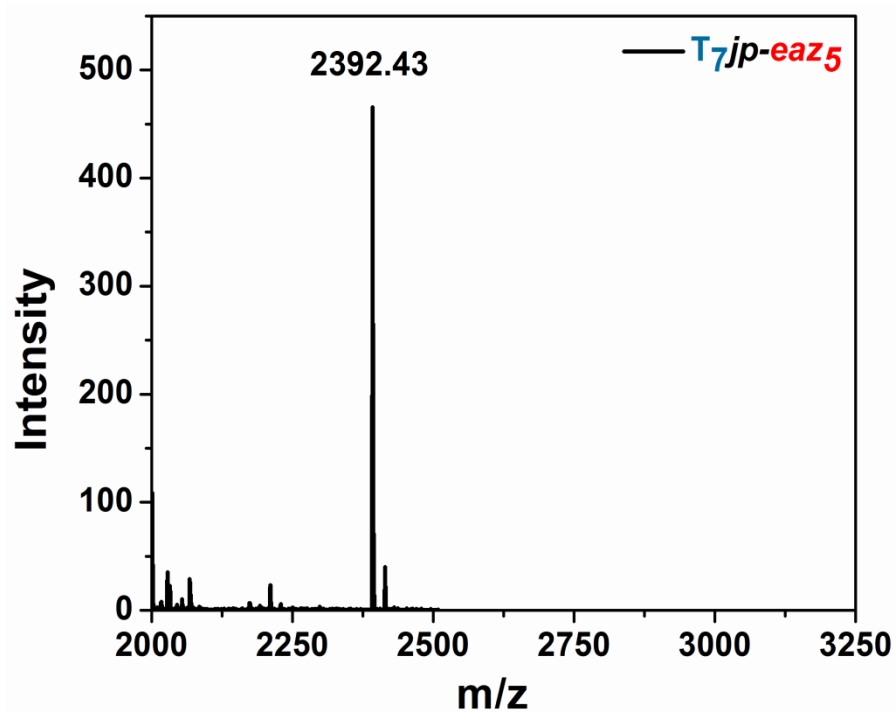
MALDI-TOF Spectrum of (PNA 9)
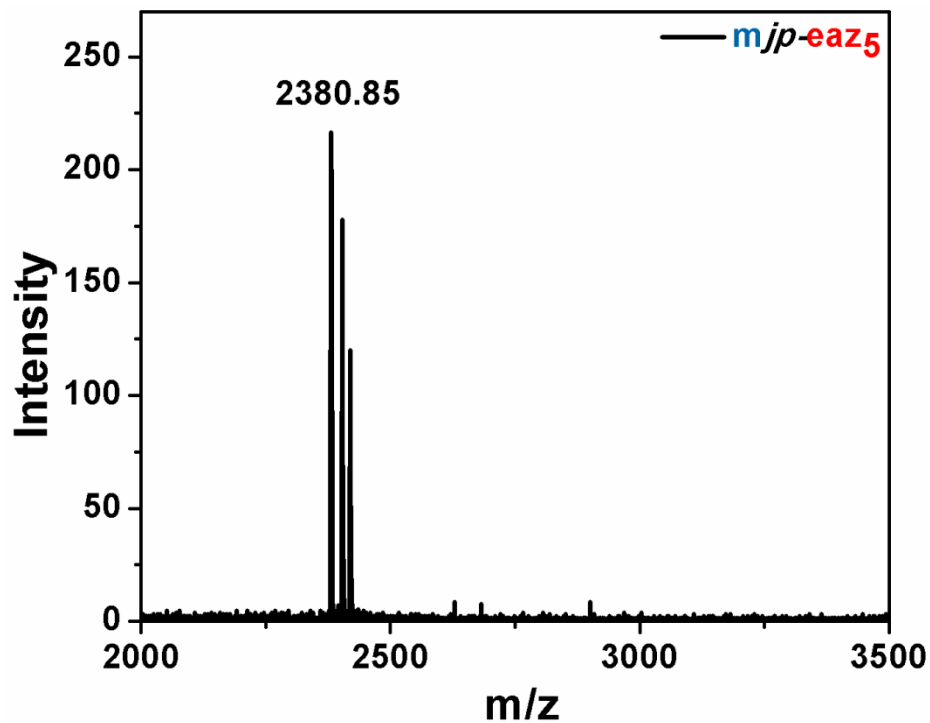
MALDI-TOF Spectrum of (JP 10)
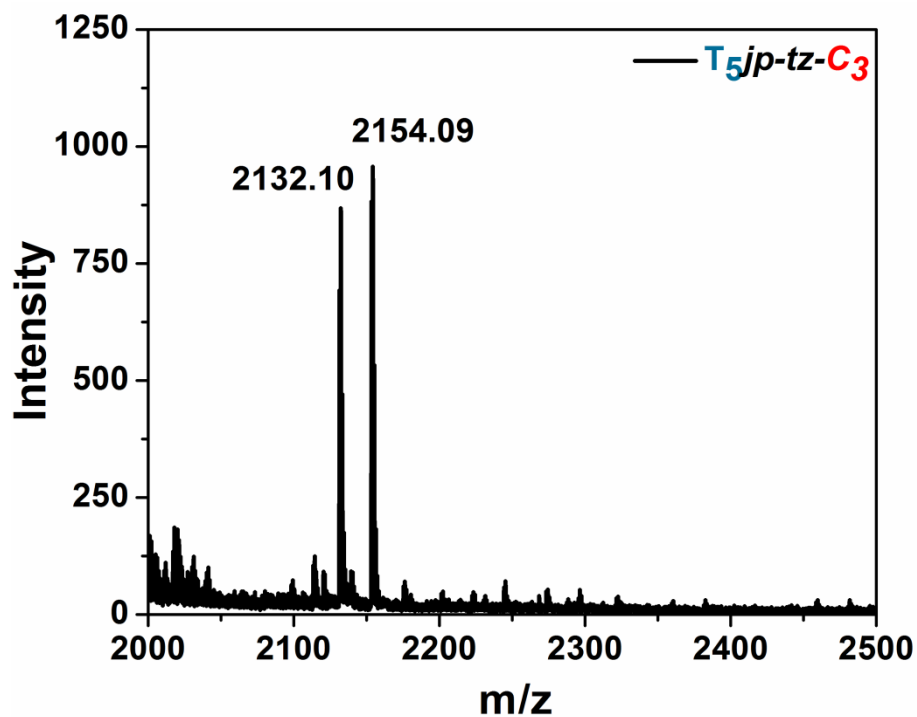
MALDI-TOF Spectrum of (JP 11)
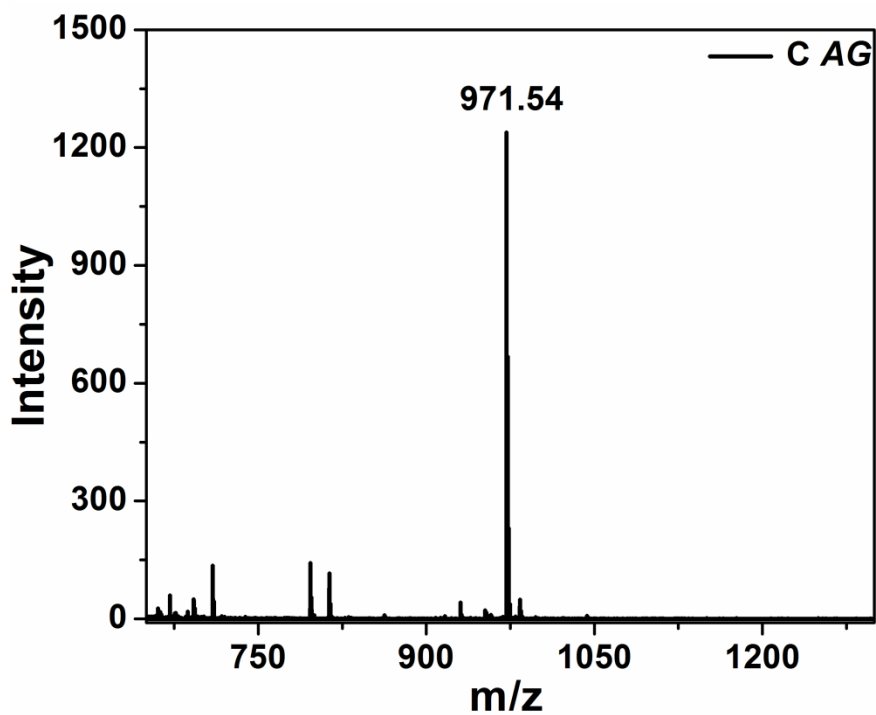
MALDI-TOF Spectrum of (JP 12)
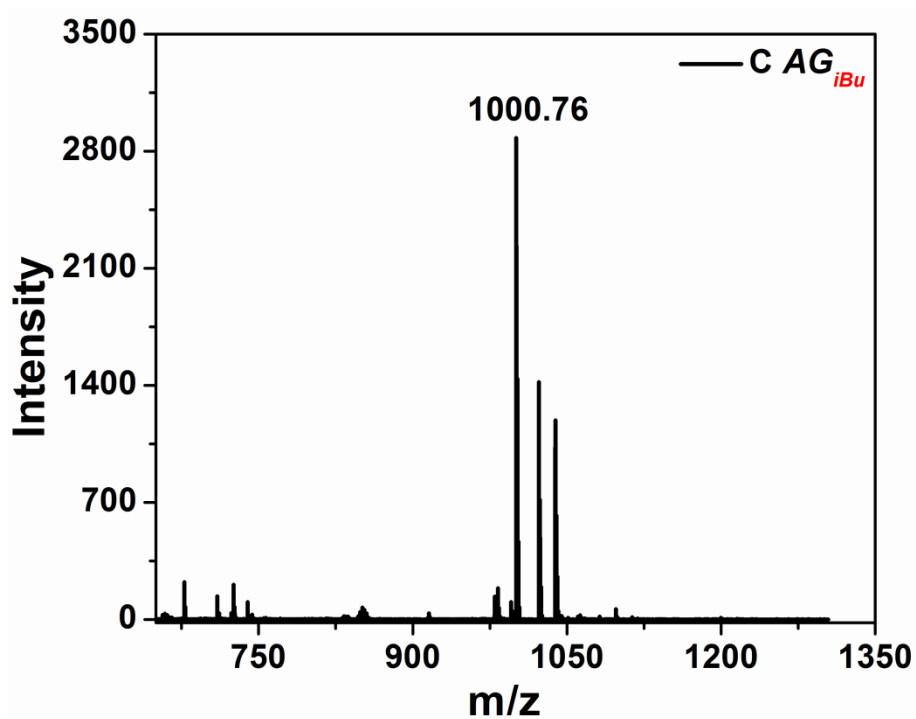
MALDI-TOF Spectrum of (PNA 13)
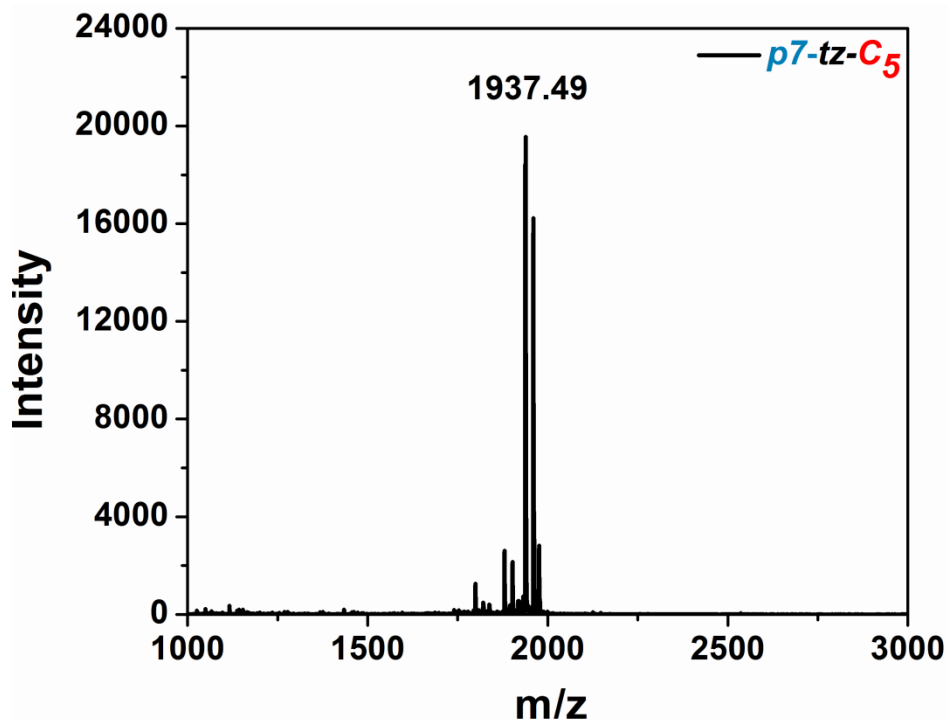
MALDI-TOF Spectrum of (PNA 14)
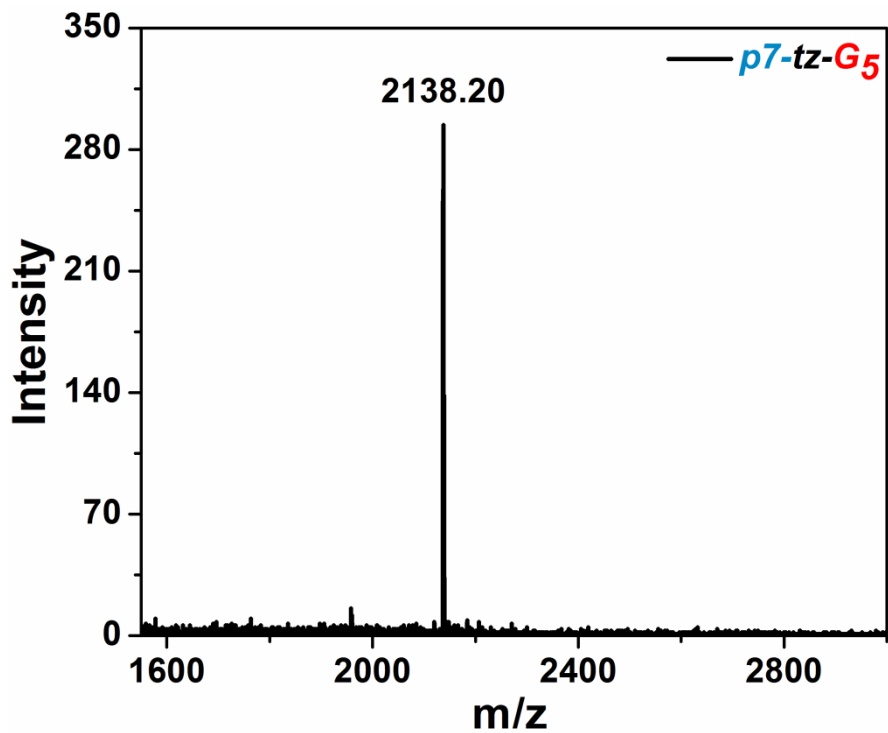
MALDI-TOF Spectrum of (PNA 15)
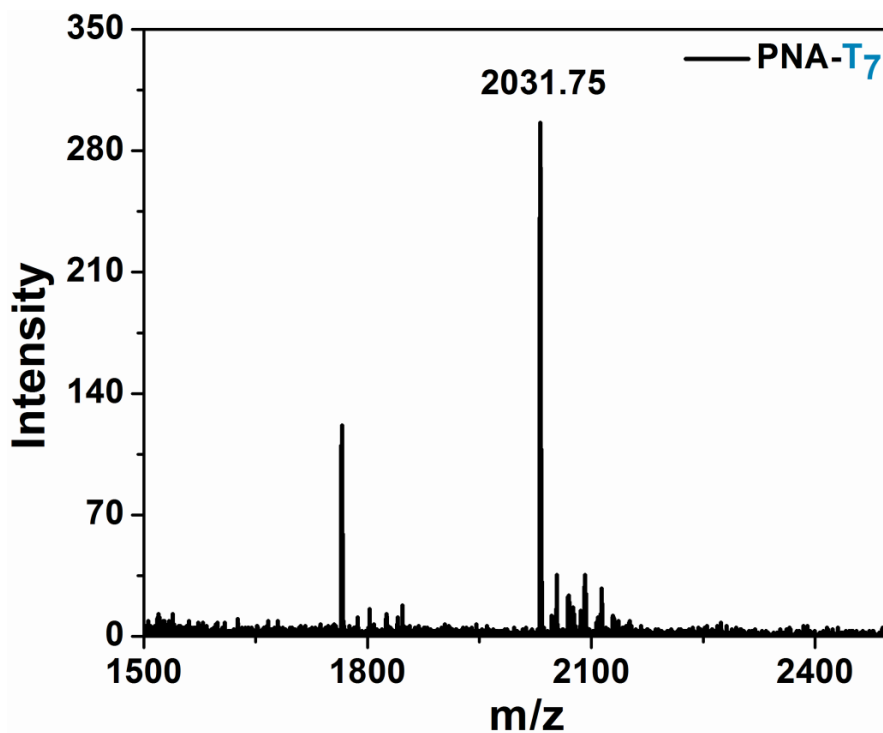
MALDI-TOF Spectrum of (**PNA 16**)
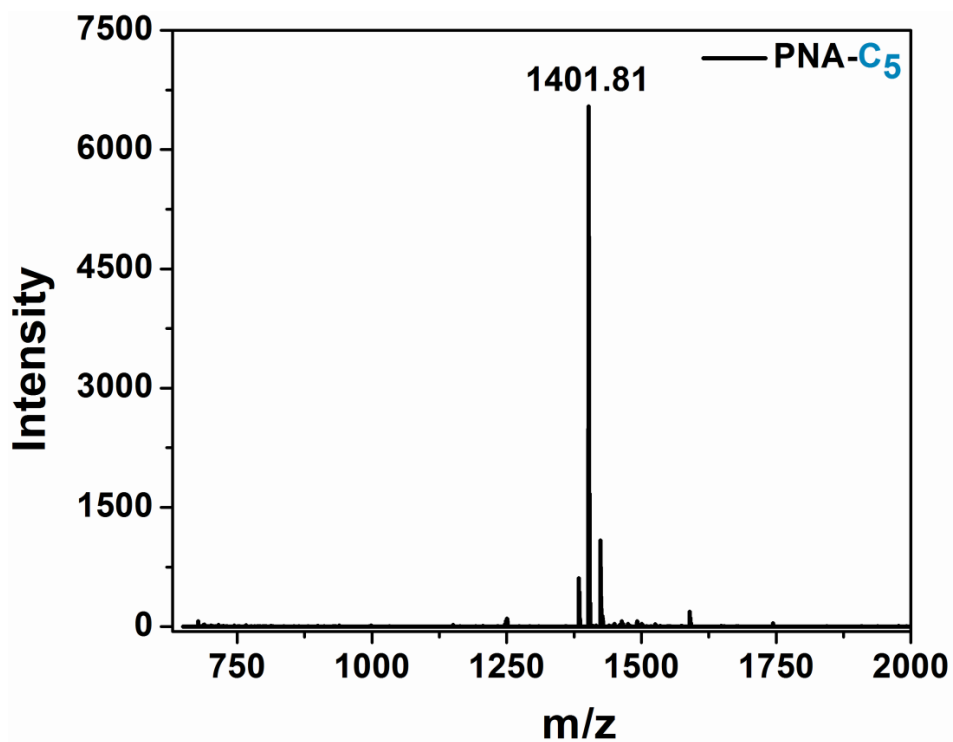
MALDI-TOF Spectrum of (PNA 17)
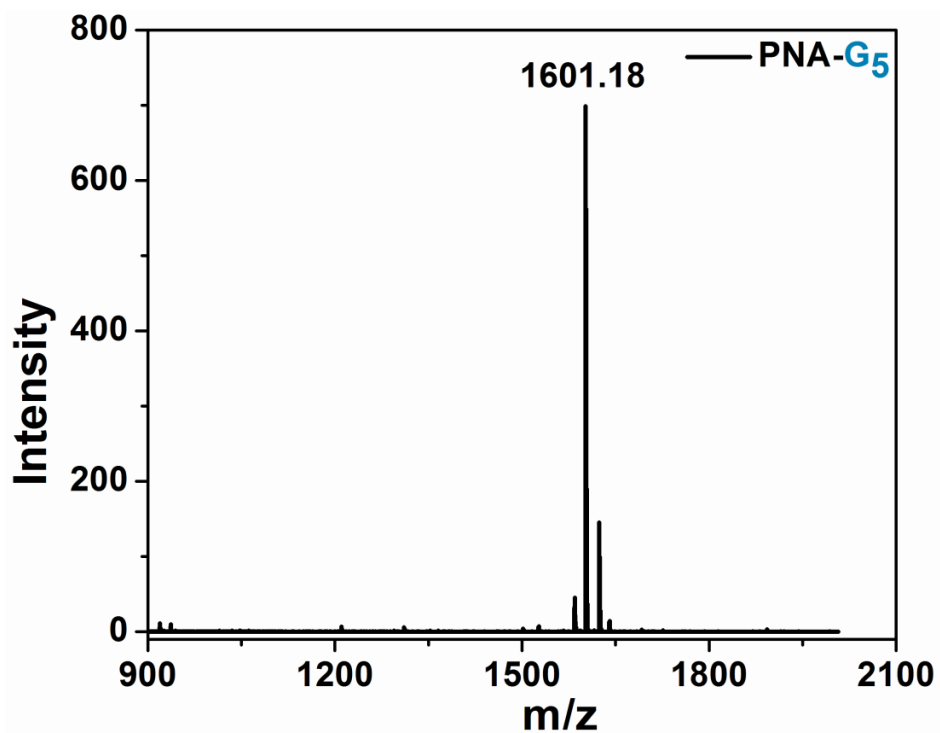
MALDI-TOF Spectrum of (PNA 18)
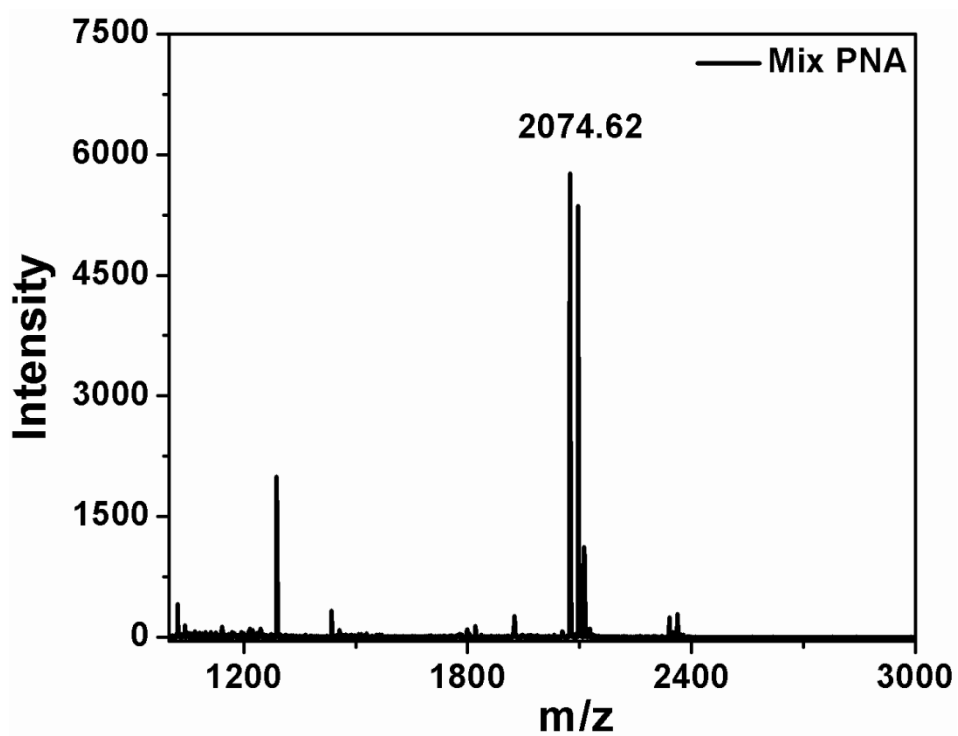
Chapter 4

Biophysical Studies of *Janus* PNA Oligomers

4.0 Introduction

The ability of modified oligonucleotides to bind to the target complementary DNA/RNA can be investigated *in vitro* by employing various biophysical techniques such as UV Job plot (stoichiometry), temperature-dependent UV-melting (UV-T_m), Circular Dichroism (CD) [1] for conformational features adopted by duplexes,[2] Isothermal Titration Calorimetry (ITC) [3] for evaluating thermodynamic features of binding and ESI-MS for characterization. [4-7]

The preceding chapter discussed the synthesis of rationally designed C_α - substituted PNA analogs. The modified PNA monomers were introduced by solid phase synthesis to generate various modified PNA and *Janus* PNA oligomers to study the effect of chirality and side chain functional groups in influencing the binding properties to complementary oligonucleotides. This chapter reports comparative biophysical studies of modified PNA and *Janus* PNA oligomers and their hybrids with complementary nucleic acids using temperature dependent UV absorbance, CD spectroscopy,[8] ITC[9,10] and ESI-MS. The objective is to understand the relative binding efficiency and selectivity of nucleobase sequences on each face (amide / triazole) of the *Janus* PNA, individually and simultaneously with complementary DNA sequences. It would provide us details of the binding strength, face selectivity and the synergistic effects of binding of cDNA on one face (*eg*. amide face) with the binding of cDNA on other face (triazole face) and *vice versa*. The homo, chimeric, hetero, and self-complementary mix *Janus* PNAs were evaluated for formation of single duplexes from each face and double duplexes from both face, in terms of binding stoichiometry by Job plot,[11] thermal stability of each of the duplexes using temperature dependent UV absorbance,[1] conformation and effects on sequential formation of double duplexes from each initial complex by CD spectroscopy. Non-covalent interaction of *Janus* PNA with DNA to form duplex, triplex or double duplex of triplex are supported by ESI-MS data and thermodynamic parameters of duplex formation by ITC.[12-14] In addition, formation of triplexes (*Janus* PNA_2:DNA), duplexes (*Janus* PNA:DNA), double duplex of triplex [DNA:*Janus* PNA:DNA:*Janus* PNA:DNA] and double duplex (DNA:*Janus* PNA:DNA) from *homo* and *chimeric Janus* PNAs are also explored.

4.1 Objectives of the present work
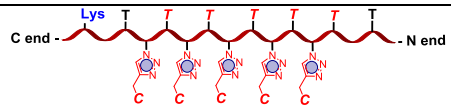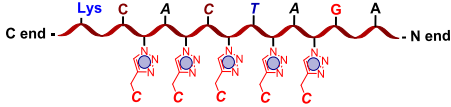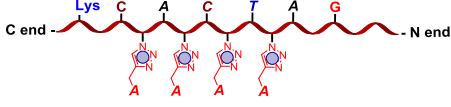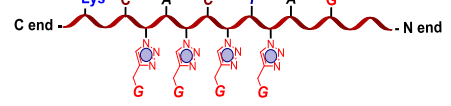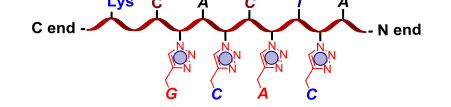
The work presented in this chapter involves biophysical studies of the *Janus* PNAs in evaluating the following aspects:

➤ Stoichiometry of binding of *Janus* PNA with complementary DNA by UV Job plot

➤ Thermal stability, binding specificity and selectivity of *Janus* PNA with cDNA using temperature-dependent UV spectroscopy

➤ Conformational study of *Janus* PNA:DNA duplex structures by CD spectroscopy

➤ Mass spectral characterization of *Janus* PNA:cDNA complexes by ESI-MS

➤ Measurement of thermodynamic parameters of *Janus* PNA:DNA interaction by ITC

➤ Analysis and comparison of the results of various biophysical studies

4.2 *Janus* PNA (JP) oligomers used for biophysical studies

The PNA monomers having C_α-substituted side chains were incorporated into PNA oligomers by solid-phase synthesis followed by click reaction[15-17] with base propynes to obtain various kinds of *Janus* PNA C_α–triazole oligomers (Table 4.1) for biophysical studies with complementary antiparallel DNAs (Table 4.2)

Table 4.1 *Janus*-PNA oligomers

Entry	Sequence Code	PNA Sequences	*Janus* oligomers
1	T_7*jp-tz-*C_5 JP 1		*Homo Janus* PNA
2	m*jp-tz-*C_5 JP 2		
3	m*jp-tz-*A_4 JP 3		*Chimeric Janus* PNA
4	m*jp-tz-*G_4 JP 4		
5	m*jp-tz*$_4$*-CACG* JP 5		*Hetero Janus* PNA

6	m*jp-tz*$_5$-*CCACG* JP 6	*Hetero Janus* PNA
7	*SCM*-*JP* JP 7	*Self Complementary Mix Janus* PNA
8	T$_7$*jp-eaz*$_5$ PNA 8	
9	m*jp-eaz*$_5$ PNA 9	
10	T$_5$*jp-tz*-C$_3$ JP 10	
11	C*A*-*G* JP 11	
12	C*A*-*G*$_{iBu}$ JP 12	
13	*p7-tz*-C$_5$ PNA 13	*aeg-tz* PNA
14	*p7-tz*-G$_5$ PNA 14	
15	PNA-T$_7$ PNA 15	
16	PNA-C$_5$ PNA 16	*aeg*-PNA
17	PNA-C$_5$ PNA 17	
18	Mix PNA PNA 18	

4.3 DNA oligonucleotides used for biophysical studies

The complementary DNA oligonucleotides **1-10** and mismatch DNA **11-17** (Table 4.1) that form antiparallel duplexes from both amide and triazole face were obtained from

commercial sources and used without further purification in the biophysical studies of PNAs or *Janus* PNAs.

Table 4.2 Complementary (**1 - 10**) and mismatch (**11 – 17**) **DNA** used for biophysical studies of *Janus* PNA

Entry	DNA	Sequence (5' to 3')
1	DNA **1**	AAA AAA AA or **dA$_8$**
2	DNA **2**	GGG GGG or **dG$_6$**
3	DNA **3**	GTG ATC T or **cDNA**
4	DNA **4**	CGT GGA
5	DNA **5**	GTG ATC TTT TGT GC or DNA **5$_{hp}$**
6	DNA **6**	GTG ATA GCT
7	DNA **7**	GTG ATA GCT TTT GGT GC or DNA **7$_{hp}$**
8	DNA **8**	GAG GAG AC
9	DNA **9**	TCC TCT
10	DNA **10**	CCC CCC or **dC$_6$**
11	DNA **11m**	AAA CAA AA
12	DNA **12m**	GGT GGG
13	DNA **13m**	GTA GTC T
14	DNA **14m**	GGT TGG
15	DNA **15m**	GTG CTA GCT
16	DNA **16m**	CGC GGA
17	DNA **17m**	CCT CCC

hp = hairpin, m = mismatch; all DNAs were designed to bind in antiparallel manner

Various *Janus* PNA (Figure 4.1) sequences were hybridized with appropriate complementary DNA (Table 4.1) in antiparallel orientation to generate PNA:DNA complexes that involved triplex, duplex, double duplex and duplex of triplex as shown in Figures 4.2 - 4.7

JP 1:DNA1:JP 1

PNA₂:DNA homo triplex

A

JP 1:DNA 2

Triazole face homo duplex

B

DNA 2:JP 1:DNA 1:JP 1:DNA 2

Double triazole face duplex of homo amide face triplex

C
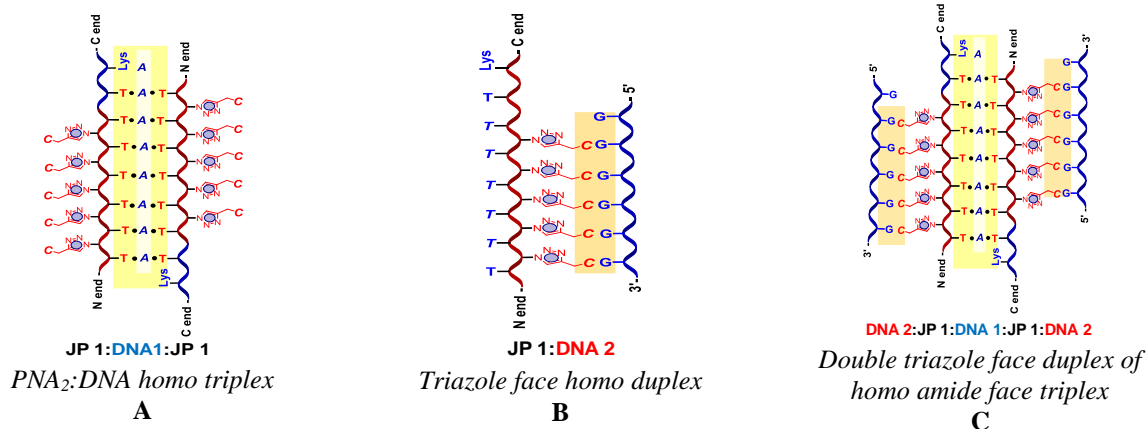
Figure 4.2 *Homo Janus* PNA:DNA complex **(A)** Triplex C_5-*tz-jp*T_7:*dA₈*:T_7*jp-tz*-C_5 (**JP 1:DNA 1:JP 1**) **(B)** Triazole face duplex T_7*jp-tz*-C_5:*dG₆* (**JP 1:**DNA **2**) **(C)** Duplex of a triplex, *dG₆:C₅*-*tz-jp*T_7:*dA₈*:T_7*jp-tz*-C_5:*dG₆* (DNA **2:JP 1:DNA 1:JP 1:DNA 2**)

DNA 3:JP 2

Hetero amide face duplex

A

JP 2:DNA 2

Homo triazole face duplex

B

DNA 3:JP 2:DNA 2

Hetero amide, homo triazole double duplex

C
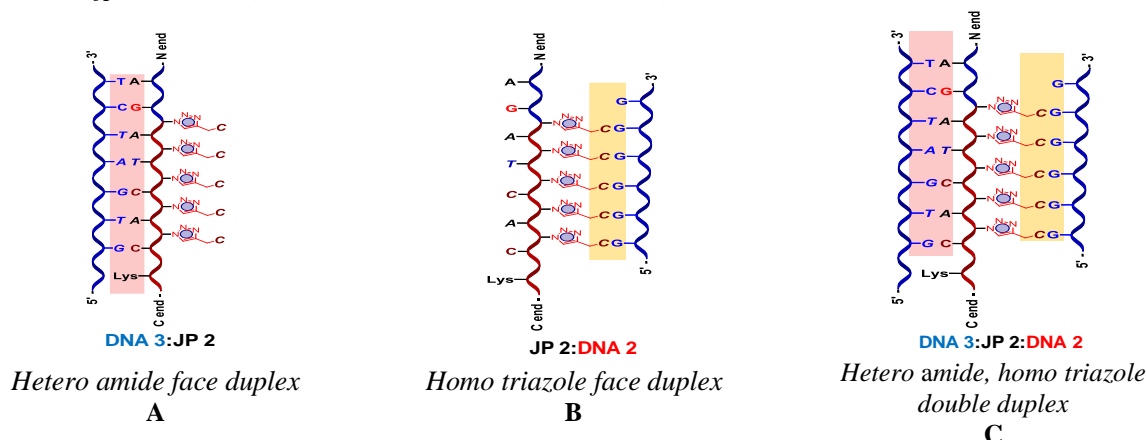
Figure 4.3 *Chimeric Janus* PNA:DNA antiparallel complexes **(A)** Hetero amide face duplex, cDNA:m*jp-tz*-C_5 (DNA **3:JP 2**) **(B)** Homo triazole face duplex m*jp-tz*-C_5:dG₆ (**JP 2:DNA 2**) **(C)** *Hetero amide, homo triazole double duplex* cDNA:m*jp-tz*-C_5:dG₆ (DNA **3:JP 2:DNA 2**)

DNA 3:JP 5

Hetero amide face duplex

A

JP 5:DNA 4

Hetero triazole face duplex

B

DNA 3:JP 5:DNA 4

Hetero amide-triazole double duplex

C

DNA 5ₕₚ:JP 5:DNA 5ₕₚ

Hairpin duplex

D
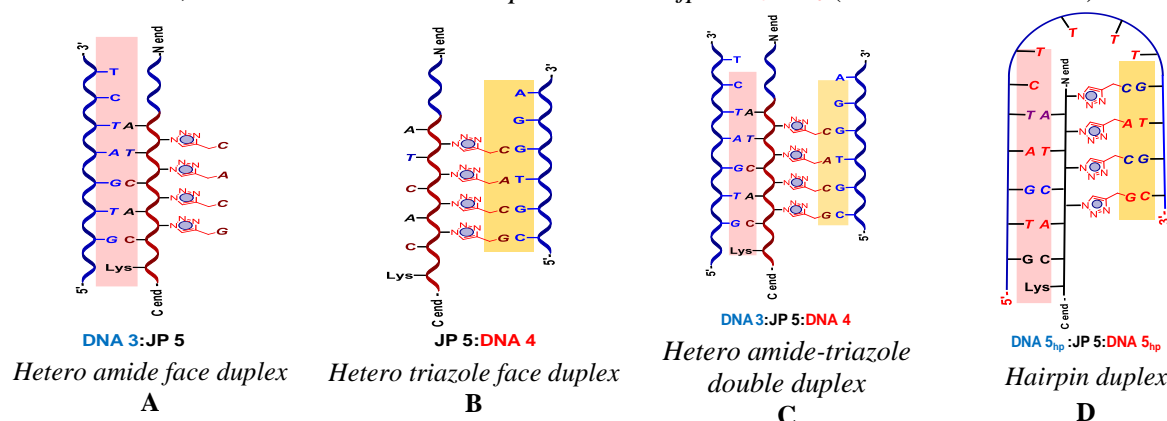
Figure 4.4 *Hetero Janus* PNA:DNA antiparallel complex **(A)** Hetero amide face duplex , DNA **3**:m*jp-tz₄-CACG* (**DNA 3:JP 5**) **(B)** *Hetero triazole face duplex,* m*jp-tz₄-CACG*:DNA **4** (**JP 5:DNA 4**) **(C)** *Hetero amide-triazole double duplex* DNA **3**:m*jp-tz₄-CACG*:DNA **4** (DNA **3:JP 5:DNA 4**) **(D)** *Hairpin duplex*DNA **5**ₕₚ:m*jp-tz₄-CACG:*DNA **5**ₕₚ (DNA **5ₕₚ:JP 5:DNA 5ₕₚ**)

DNA 6:JP 6
Amide face hetero duplex
A

JP 6:DNA 4
Triazole face hetero duplex
B

DNA 6:JP 6:DNA 4
*Amide-Triazole face
hetero double duplex*
C

DNA 7hp:JP 6:DNA 7hp
Hairpin duplex
D
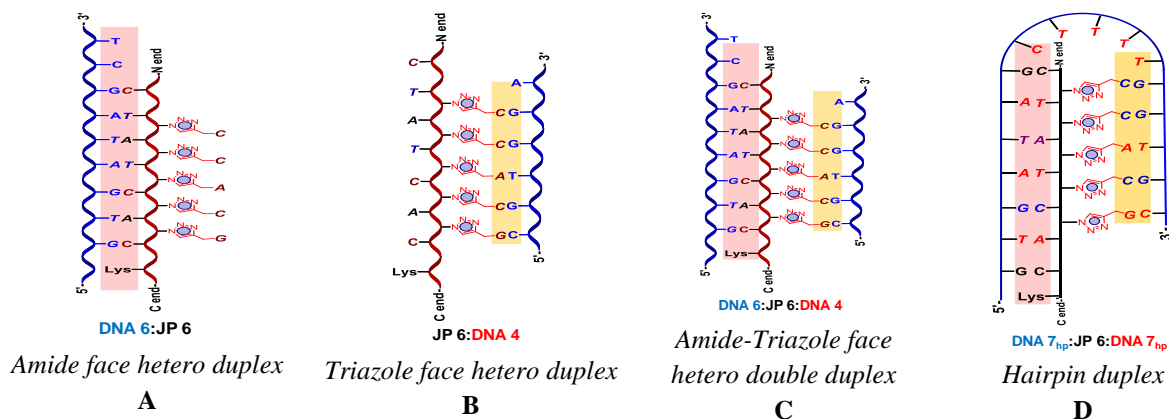
Figure 4.5 *Hetero Janus* PNA:DNA antiparallel complex **(A)** hetero amide duplex DNA **6:JP 6** or DNA **6**:m*jp-tz5-CCACG* **(B)** hetero triazole duplex **JP 6:DNA 4** or m*jp-tz5-CCACG*:DNA **4 (C)** hetero double duplex DNA **6:JP 6:DNA 4** or DNA 6:m*jp-tz5-CCACG*:DNA **4 (D)** hetero hairpin double duplex DNA **7hp:JP 5:DNA 5hp** or DNA **7hp**:m*jp-tz5-CCACG:*DNA **2hp**

DNA 8:JP 7
Amide face heteroduplex
A

JP 7:DNA 9
Triazole face heteroduplex
B

DNA 8:JP 7:DNA 9
Hetero double duplex
C
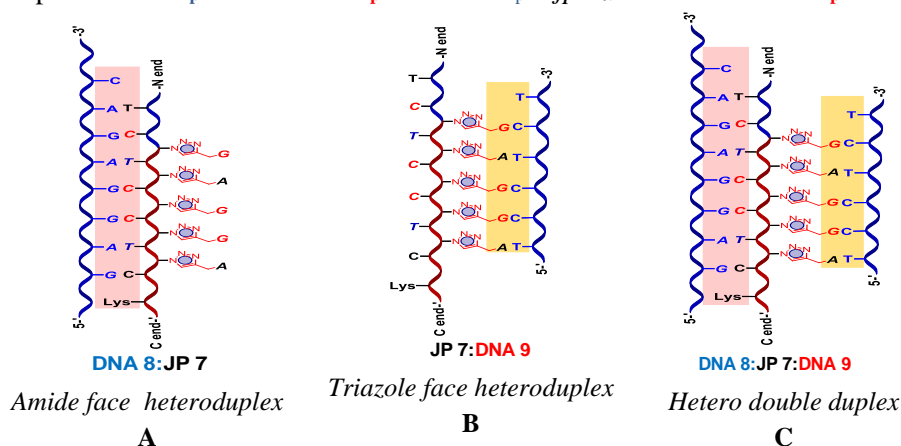
Figure 4.6 *Self-complementary Mix Janus* PNA(*SCM-JP*):DNA antiparallel complex **(A)** DNA **8**:*SCM-JP* (DNA **8:JP 7**) **(B)** *SCM-JP:*DNA **9** (**JP 7:DNA 9**) **(C)** DNA **8**:*SCM-JP:*DNA **9** (DNA **8:JP 7:DNA 9**)

PNA 13:DNA 2
A

PNA 14:DNA 10
B

PNA 16:DNA 2
C
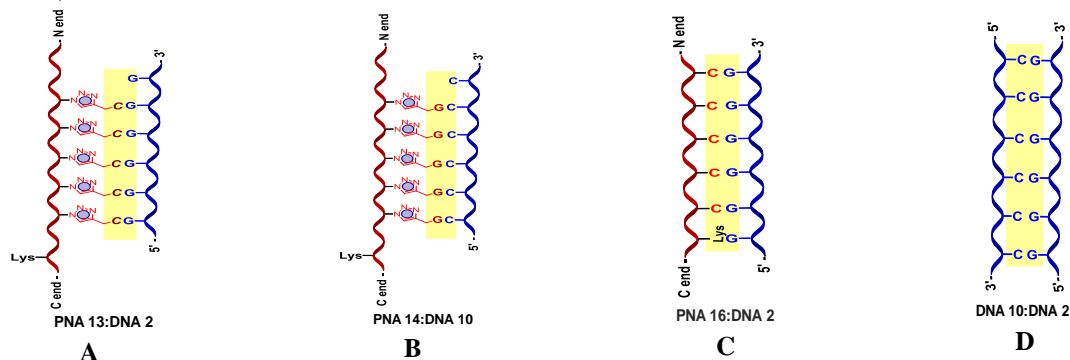
DNA 10:DNA 2
D

Figure 4.7 Antiparallel duplex for control studies **(A)** *p7-tz-C5*:dG6 (PNA **13:**DNA **2**) **(B)** *p7-tz-G5*:dC6 (PNA **14:**DNA **10**) **(C)** PNA C5:dG6 (PNA **16:**DNA **2**) **(D)** dC6:dG6 (DNA **10:**DNA **2**)

4.4 Results

The interaction of *Janus* PNAs with complementary DNA for complexation was investigated by changes in UV, CD, UV-Job plot, ESI-MS, and ITC. The results are presented in the following six sections:

Section 4.4.1: Binding studies with *non-Janus **aeg-tz**-PNA* with complementary DNA

Section 4.4.2: *Homo Janus* PNA (T_7*jp-tz-C₅* or **JP 1**)

Section 4.4.3: *Chimeric Janus* PNA (m*jp-tz-C₅* or **JP 2**)

Section 4.4.4: *Hetero Janus* PNA (m*jp-tz₄-CACG* or **JP 5** and m*jp-tz₅-CCACG* or **JP 6**)

Section 4.4.5: *Hetero Janus* PNA *hairpin* DNA duplexes

Section 4.4.6: *Self-Complementary Mix Janus* PNA (*SCM-JP* or **JP 7**)

4.4.1 Binding studies with *non-Janus tz*-PNA with complementary DNA

In this section the UV-T_m experiments of DNA duplexes from *non-Janus triazole* PNA (*aeg-tz* PNA) and amide *aeg*-PNA with the complementary DNAs are reported to compare the individual roles of amide-nucleobases and triazole-nucleobases in stabilizing the PNA:DNA duplexes.

4.4.1a Thermal stability of DNA / *aeg* PNA / *aeg-tz* PNA (*p7-tz-C₅*) with complementary DNA duplexes

The *Janus* PNA surrogate *p7-tz-C₅* PNA (*p7-tz-C₅*) containing nucleobases linked to backbone at C_α with *S*-ethyltriazole linker instead of N-tertiary amide substituted nucleobases is a new structure hitherto unknown and its ability to bind to cDNA needs to be first established. Hence hybridization studies were carried out with *p7-tz-C₅* (PNA **13**) that has only triazole linked nucleobases and compared with that of standard *aeg*-PNA-C_5 (PNA **16**). The latter (PNA **16**) forms hybrid PNA-**C₅:**dG_6 (DNA **2**) only from amide face, while the triazole PNA *p7-tz-C₅* (PNA **13**) should form duplex *p7-tz-C₅:*dG_6 with dG_6 (DNA **2**) only from triazole face, without the effects of N-tertiary amide linkage. The PNA *p7-tz-C₅* formed duplex *p7-tz-C₅:*dG_6 (PNA **13:**DNA **2**) showing nice sigmoidal transition with single T_m of 40.8 ℃ (Figure 4.8). This duplex is as stable as the duplex from *aeg*-PNA **16:**dG_6 (DNA **2**) with T_m of 41.8 °C and DNA:DNA duplex (dC_6:dG_6) with T_m of 40 ℃ (Table 4.2). Thus, it was significant to find that the new backbone PNA analogue triazole *p7-tz-C₅* (PNA **13**) effectively formed duplex with

cDNA, as good as the well-studied standard amide *aeg* PNA **16** in forming duplex with similar thermal stability.
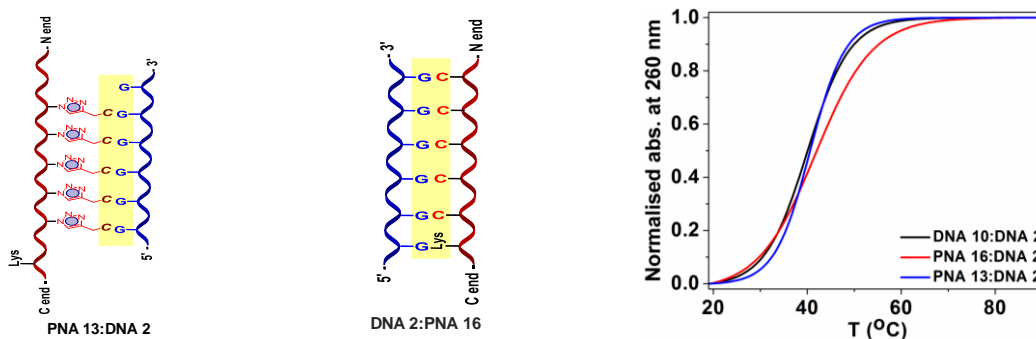
Figure 4.8 Temperature dependent UV absorbance curves for DNA duplexes of *aeg* PNA/*aeg-tz* PNA/DNA. PNA **13** = *p*7-*tz*-C_5, PNA **16** = PNA-C_5, DNA **2** = 5'-GGGGGG-3'; Buffer: 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM

Table 4.2 UV-T_m (ºC) values of DNA duplexes with *aeg/aeg-tz* PNA

Entry	PNA:DNA complexes	T_m (ºC)		ΔT_m (ºC), duplex
		Amide	Triazole	(PNA:DNA) - (DNA:DNA)
1	PNA **13**:DNA **2** (*p*7-*tz*-C_5:dG_6)	-	40.8	+0.8
2	DNA **2**:PNA **16** (PNA-C_5:dG_6)	41.8	-	+1.8
3	PNA **14**:DNA **10** (*p*7-*tz*-G_5:dC_5)	63.9		+23.9
4	PNA **17**:DNA **10** (PNA-G_5:dC_6)	63.2		+23.2
5	DNA **2**:DNA **10** (dC_6:dG_6)	40.0		-

Buffer, 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM T_m values are accurate to ±1.0 ºC

4.4.1b Thermal stability of *aeg* PNA/ *aeg-tz* PNA (*p*7-*tz*-G_5) with complementary DNA duplexes.

In order to examine the generality of the role of triazole linked nucleobases to form PNA duplex and sequence effects on stabilization of duplexes, the triazole PNA *aeg-tz* *p*7-*tz*-G_5 (PNA **14**) in which G_5 is linked to triazole side chain was duplexed with complementary dC_6 (DNA **10**). This also exhibited a nice sigmoidal transition with T_m of 63.9 ºC, as good as that of the corresponding standard *aeg*-PNA-G_5 that formed duplex PNA-G_5:dC_6 (PNA **17**:DNA **10**) with a T_m of 63.2 ºC (Figure 4.9). Both these T_ms are considerably higher than the T_m of 40 ºC for DNA:DNA duplex dG_6:dC_6. This further established that the new triazole PNA analogue (*aeg-tz*) where the bases are linked through triazole form duplexes with complementary DNA,

as efficient as the well-studied standard *aeg* PNA. The *aeg-tz* PNA is a new PNA analog as good as standard N-tertiary amide PNA.
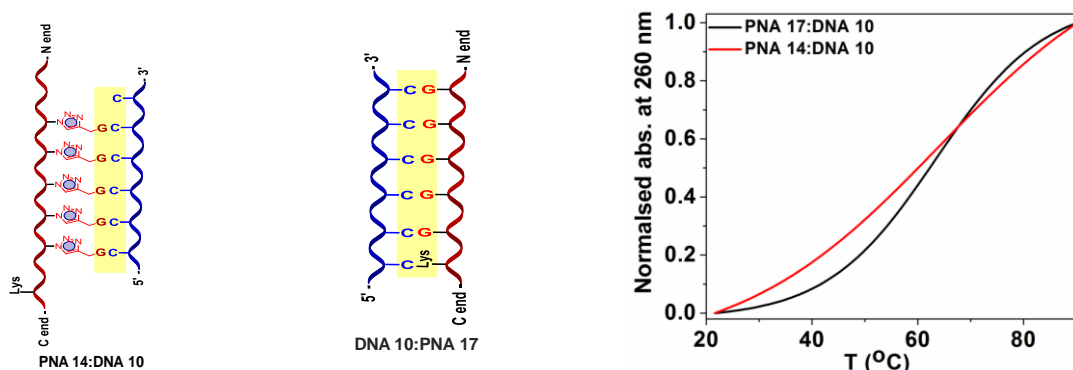
Figure 4.9 Temperature dependent UV absorbance curves for complementary DNA/ *aeg* PNA/ *aeg -tz* PNA:DNA duplexes, PNA **14** = *p*7-*tz*-G_5, PNA **17** = PNA-G_5 and DNA **10** (dC$_6$= 5' CCCCCC 3') Buffer: 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM.

To establish the sequence specificity of base pairing in triazole *p*7-*tz*-C_5/G_5 PNA:DNA duplexes, melting studies were done using DNA **12m** (5'-GG**T**GGG-3') that has single T-C mismatch while pairing with *p*7-*tz*-C_5 (PNA **13**) and with DNA **17m**) (5'-CC**T**CCC-3') that has single T-G mismatch while pairing with *p*7-*tz*-G_5 (PNA **13**) (Figure 4.10). The mismatched duplex *p*7-*tz*-C_5 (PNA **13**):DNA **12m** showed a T_m of 40.3 ºC, lower by 0.5 ºC (Table 4.3, entry 1) compared to T_m of 40.8 ºC in perfect match duplex with DNA **2**. The mismatch duplex of *aeg*-PNA **16**:DNA **12m** showed T_m of 35.2 ºC that is 6.6 ºC lower (Table 4.3, entry 2) compared to T_m (41.8 ºC) of perfect match duplex with DNA **2**. The duplex *p*7-*tz*-G_5 PNA **14**:DNA **17m** had a T_m of 55.6 ºC (Table 4.3, entry 3) that was 8.3 ºC lower compared to T_m of 63.9 ºC of its perfect duplex with DNA **10**. The T_m of corresponding *aeg*-PNA **17**:DNA **17m** duplex was 61.2 ºC (Table 4.3, entry 4) which is 2.0 ºC less compared to T_m of 63.2 ºC for perfect match duplex with DNA **10**. The lowering of T_m in mismatched duplexes substantiate the formation of sequence specific binding of duplexes of *aeg-tz* PNA by complementary base pairing.
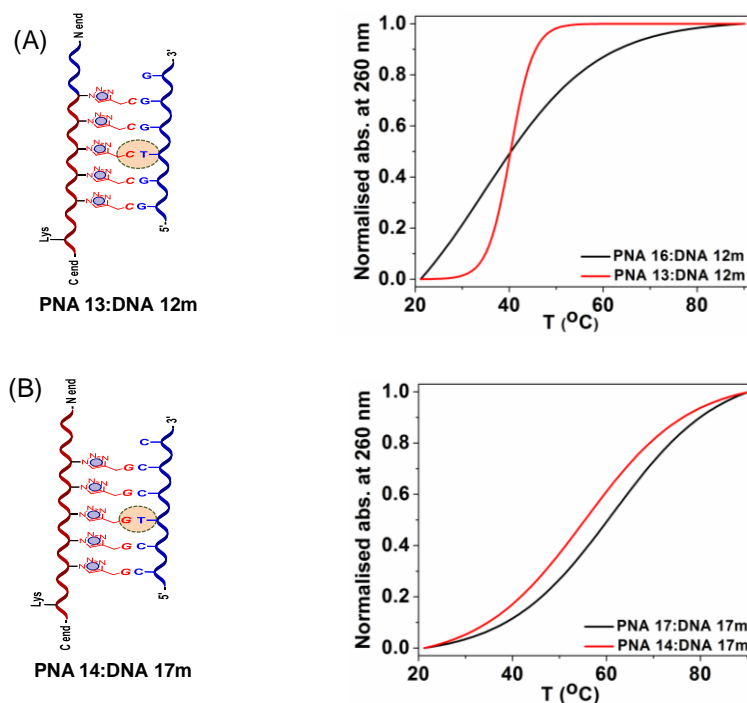
Figure 4.10 Temperature dependent UV absorbance curves for complementary DNA/ *aeg* PNA/ *aeg -tz* PNA:DNA duplexes, PNA 13 = *p7-tz-C₅,* PNA **14** = *p7-tz-G₅,* PNA 16 = PNA-C₅, PNA **17** = PNA-G₅ and DNA **12m** = 5'-GGTGGG-3', DNA **17m** (dC₆= 5'-CCTCCC-3') Buffer: 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM, T_m values are accurate to ±1.0 °C.

Table 4.3 UV-T_m (°C) values of mismatch DNA duplexes with *aeg/aeg-tz* PNA

Entry	PNA:DNA mismatch duplexes	T_m (°C)		ΔT_m^1 (°C) PNA:DNA (mismatch) - PNA:DNA (complementary)
		Amide	Triazole	
1	*p7-tz-C₅* (PNA **13**):DNA **12m**	-	40.3	- 0.5
2	PNA-C5 (PNA **16**):DNA **12m**	35.2	-	- 6.6
3	*p7-tz-G₅* (PNA **14**):DNA **17m**		55.6	- 8.3
4	PNA-G₅ (PNA **17**):DNA **17m**	61.2		- 2.0

ΔT_m^1 indicates the difference in T_m of PNA with mismatch DNA and cDNA. m = mismatch DNA

4.4.1c CD spectra of DNA/*aeg* PNA/ *aeg-tz* PNA (*p7-tz-G₅*):DNA duplexes

The conformation of PNA:DNA, PNA:RNA and PNA:PNA duplexes have been well studied using CD spectroscopy. The effect of chiral α-side chain substitution as in triazole PNA *aeg-tz* PNA on the conformation of derived *aeg-tz* PNA:DNA duplexes was studied by CD spectroscopy.

CD spectra of p7-tz-C₅/aeg-PNA-C₅/DNA:DNA duplexes. CD spectra of the control *aeg-tz* duplexes $p7$-*tz*-C_5:dG$_6$ (*aeg-tz*-PNA **13**:DNA **2**), dC$_6$:dG$_6$ (DNA **10**:DNA **2**) and PNA-C$_5$:dG$_6$ (PNA **16**:DNA **2**) shown in (Figure 4.11). All duplexes showed a positive band around 260 nm ($\pm$ 2 nm) and a negative band around 240 nm ($\pm$ 2 nm) with cross-over points at 247 nm. The CD profile of the duplex from triazole PNA $p7$-*tz*-C_5:dG$_6$ is very similar, to that of well characterized *aeg*-PNA and DNA duplexes, suggesting that the new backbone PNA $p7$-*tz*-C_5 is equally effective in forming duplex with DNA as much as the standard *aeg*-PNA.
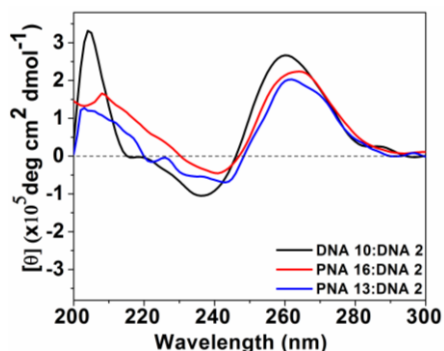
Figure 4.11 CD spectra of duplexes PNA **13** = $p7$-*tz*-C_5, PNA **16 =** PNA-**C₅,** DNA **2** (dC$_6$= 5' GGGGGG 3'), and DNA **10** (dC$_6$= 5' CCCCCC 3'), Buffer: 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM.

The combined UV and CD results of Section 4.2.1 (well defined sigmoidal transition in temperature dependent absorbance, lower T_m in mismatched DNA duplexes and CD profile (similar to that of PNA:DNA duplexes) suggest that the triazole PNAs $p7$-*tz*-C_5/G_5 in which the nucleobases are linked at C_α of aminoethylglycyl (*aeg*) backbone through a linker containing triazole moiety form sequence specific duplexes via complementary base pairing. These duplexes are as stable as that of corresponding *aeg* PNA in which nucleobases are attached to backbone via tertiary amide linkage. Further, G bases linked to triazole PNA $p7$-*tz*-G_5 (PNA **14**) bind with higher stability to complementary DNA dC$_6$, as compared to inverse situation of triazole PNA $p7$-*tz*-C_5 (PNA **13**) binding to complementary dG$_6$. This trend is similar to that seen with *aeg*-PNA-C$_6$ (PNA **16)** and *aeg*-PNA-G$_6$ (PNA **17**).

After first establishing the stand-alone base pairing properties of triazole PNAs, the next sections report on studies with Janus PNAs which are composite molecules in which nucleobases are attached on one face by tertiary amide linkage and on the other face through triazole linkage. The Janus PNAs are thus endowed with structures that enable them base pair simultaneously with two corresponding complementary DNAs.

4.4.2 Binding studies of *Homo Janus* PNA (T7*jp-tz-C5*) with cDNA

4.4.2a UV Job plot study of *homo Janus* duplex T7*jp-tz-C5* (JP 1):dG6 (DNA 2)

Determination of the binding stoichiometry in *Janus* PNA:DNA duplexes is crucial for constituting their duplexes and determination of their thermal stabilities. The stoichiometry of duplexes was established by Job plot monitored by UV absorbance of different mole fractions. The discontinuation or an intersection point corresponds to the binding stoichiometry in PNA:DNA complexes. The binding stoichiometry experiment for generating duplex alone from triazole side was done by mixing varying ratios of T7*jp-tz-C5* (JP 1, 5 μM) and oligonucleotide dG6 (DNA 2, 5 μM) keeping the total concentration constant (5 μM) in 10 mM sodium cacodylate (pH 7.1) containing 10 mM NaCl. Each sample was heated to 90 °C, slowly cooled to 20 °C (annealing) and then kept in the refrigerator at 3 °C for 8 h. The ratio of UV absorbance (260 / 290 nm) of T7*jp-tz-C5*:dG6 (JP 1:DNA 2) was subtracted with that of control PNA-T7 (260 / 290 nm) at each point and plotted against relative molar ratio (Figure 4.12). The break point in the data indicated stoichiometry 1:1 corresponding to formation of duplex from triazole face. The stoichiometry for other *Janus* PNA:DNA duplexes were derived from ITC plots as described detail in the section 4.4.2e
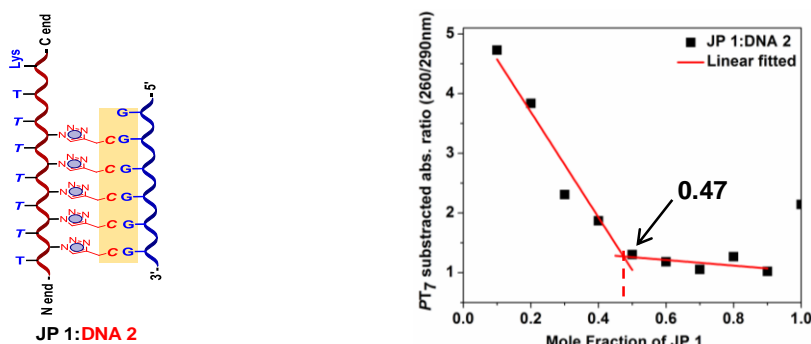
Figure 4.12 Continuous variation Job plot of homo triazole duplex T7*jp-tz-C5*:dG6 (JP 1:DNA 2) showing stoichiometry 1:1

4.4.2b Thermal stability of *homo Janus* PNA T7*jp-tz-C5*:DNA duplexes

The hybridization studies of C_α-substituted *Janus* PNA oligomer were done individually with its complementary DNA on amide face and that on triazole face by temperature-dependent changes in UV-absorbance. The thermal stabilities of duplex, triplex and duplex of triplex were determined from UV-temperature plots. Control experiments were done for the effect of C_α-

aeg-PNA side chain alone on the thermal stabilities of derived *aeg-tz* PNA:DNA complexes (4.4.1a/b). The T_m values were obtained from the mid-points of the thermal stabilities of various modified PNAs.

Homo Janus PNA T_7jp-tz-C_5 (**JP 1**) was hybridized individually with amide face cDNA **1** (dA_8) and triazole face cDNA **2** (dG_6) and the UV-T_m plots are shown in Figure 4.13**.** The T_m values were obtained from the peak maxima of the derivative plots. The triazole duplex T_7jp-tz-C_5:dG_6 (**JP 1:DNA 2**) gave T_m of 50.7 ºC (Table 4.4, entry 1). Attempts to make the amide duplex dA_8:T_7jp-tz-C_5 resulted in the formation of the corresponding triplex C_5-tz-jpT_7:dA_8:T_7jp-tz-C_5 (**JP 1:DNA 1:JP 1**) that showed a single transition with T_m of 30.2 ºC (Table 4.4, entry 2). Thus the triazole duplex T_7jp-tz-C_5:dG_6 (**JP 1:DNA 2**) shows a higher T_m (+20.5 ºC) compared to amide triplex inspite of having less number of base pairs. The hybridization of the triplex C_5-tz-jpT_7:dA_8:T_7jp-tz-C_5 with dG_6 (DNA **2**) generated the double duplex of triplex dG_6:C_5-tz-jpT_7:dA_8:T_7jp-tz-C_5:dG_6 (DNA **2:JP 1:DNA 1:JP 1:DNA 2**) that showed two different melting temperatures T_{m1} (45.8 °C) and T_{m2} (73.5 °C) (Table 4.4, entry 3). The higher T_{m2} conforms to the two identical duplexes dG_6:C_5-tz-jpT_7 from either side of the triplex C_5-tz-jpT_7:dA_8:T_7jp-tz-C_5 and the lower melting (T_{m1}) is that of triplex C_5-tz-jpT_7:dA_8:T_7jp-tz-C_5. Both T_{m1} (triplex) and T_{m2} (duplex) are higher than the corresponding T_ms observed with individual duplexes suggesting that both duplex and triplex are concurrently stabilized when present together. The triazole duplex showed stabilization of +22.8 °C and the amide triplex stabilized by +15.6 °C over that in isolated complexes (Table 4.4, entry 3). The differential ΔT_m between triazole duplex and amide triplex in double duplex of triplex was higher +27.7 ºC compared to that of +20.5 ºC in isolated duplexes. This suggested a synergistic effect on duplex stabilization by *Janus* PNA triplex and vice versa. The identity of triplex and duplexes are substantiated by CD and ITC resulted presented later in this section.

(A)
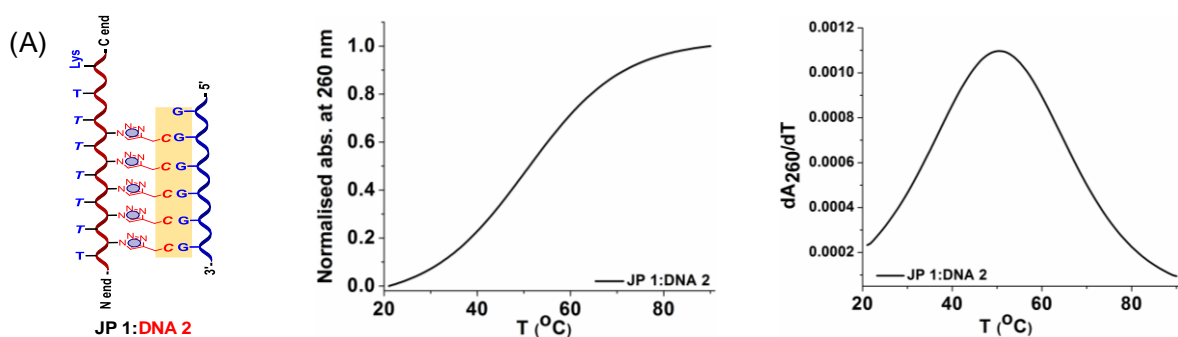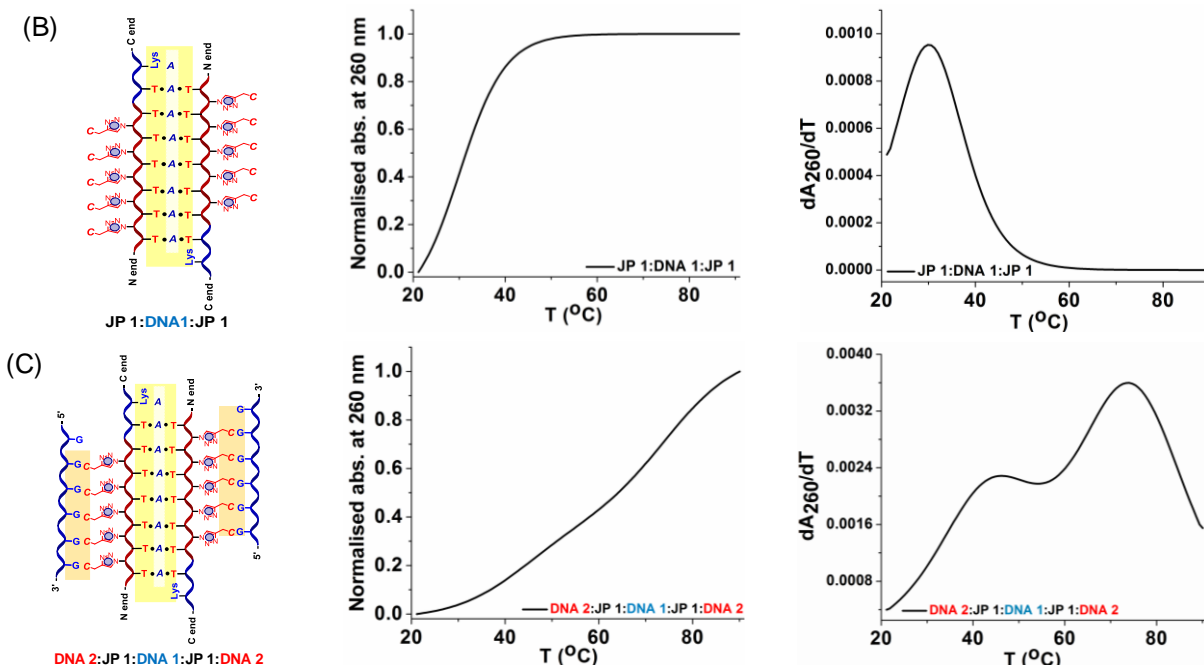
JP 1:DNA 2

Figure 4.**13** Temperature dependent UV absorbance curves for DNA complexes of T$_7$*jp-tz-C$_5$* (**JP 1**) with dA$_8$ (DNA **1**, 5'AAAAAAAA 3') and dG$_6$ (DNA **2**, 5' GGGGGG 3'); Buffer 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM **(A)** T$_7$*jp-tz-C$_5$*:dG$_6$ (**JP 1**:DNA **2**) **(B)** *C$_5$-tz-jp*T$_7$:dA$_8$:T$_7$*jp-tz-C$_5$* (**JP 1**:DNA **1**:**JP 1**) **(C)** dG$_6$:*C$_5$-tz-jp*T$_7$:dA$_8$:T$_7$*jp-tz-C$_5$*:dG$_6$ (DNA **2**:**JP 1**:DNA **1**:**JP 1**:DNA **2**), T_m values are accurate to ± 1.0 °C.

Table 4.4 UV-T_m (°C) of complexes of T$_7$*jp-tz-C$_5$* (**JP 1**) with complementary DNA **1** and DNA **2**

Entry	*Homo Janus* **PNA:DNA complexes**	T_m (°C)		ΔT_m (°C) (double duplex - T_m duplex)	
		Amide	Triazole	Amide	Triazole
1	T$_7$*jp-tz-C$_5$*:dG$_6$ (**JP 1**:DNA **2**)	-	50.7	-	-
2	*C$_5$-tz-jp*T$_7$:dA$_8$:T$_7$*jp-tz-C$_5$* (**JP1**:DNA **1**:**JP 1**)	30.2	-	-	-
3	dG$_6$:*C$_5$-tz-jp*T$_7$:dA$_8$:T$_7$*jp-tz-C$_5$*:dG$_6$ (DNA **2**:**JP1**:DNA1:**JP1**:DNA **2**)	45.8	73.5	+15.6	+22.8

Mismatched homo Janus PNA:DNA complexes. In order to understand the enhanced stability of each face duplex and attribute it to sequence specific H-bonding as per WC base pairing, melting studies were done with T$_7$*jp-tz-C$_5$* (**JP 1**) hybridized with 5'-GG**T**GGG-3' (DNA **11m**) that has a C-T mismatch on amide face and with 5'-AAA**C**AAAA-3' (DNA **12m**) that also has a C-T mismatch on triazole face (Figure 4.14). The triazole face mismatch duplex T$_7$*jp-tz-C$_5$*:DNA **12m** showed a single transition corresponding T_m = 40.8 °C (Table 4.5, entry 1), as

compared to T_m 50.7 °C observed with perfect complementary duplex T$_7$jp-tz-C$_5$:dG$_6$ (**JP 1**:DNA **2**) T_m. Thus (1xC-T) mismatch introduced destabilization of duplex by 9.9 °C. The mismatched triplex from amide face **JP 1**:DNA **11m**:**JP 1** also showed a single transition with T_m of 28.0 °C (Table 4.5, entry 2) as compared with T_m of 30.2 °C for perfect complementary triplex C$_5$-tz-jpT$_7$:dA$_8$:T$_7$jp-tz-C$_5$. Thus T_m was lower by 2.2 °C for mismatch triplex **JP 1**:DNA **11m**:**JP 1** in reference to the perfect triplex **JP1**:DNA **1**:**JP 1**.

In presence of both face mismatch DNAs (DNA **11m** and DNA **12m**), the *homo Janus* PNA C$_5$-tz-jpT$_7$ (**JP 1**) formed double duplex of triplex DNA **12m**:**JP 1**:DNA **11m**:**JP 1**:DNA **12m** and exhibited a broad melting with single transition at T_m = 53.2 °C (Table 4.5, entry 3). The observance of single transition in this double duplex of triplex is due to overlap of both duplex and triplex transitions and a lower difference in T_ms of both duplexes. The lowering of T_m of both duplex and triplex, similar to that seen in perfect duplexes supported the formation of the complex from *homo Janus* PNA C$_5$-tz-jpT$_7$ with cDNAs by sequence dependent complementary base pairing.

192
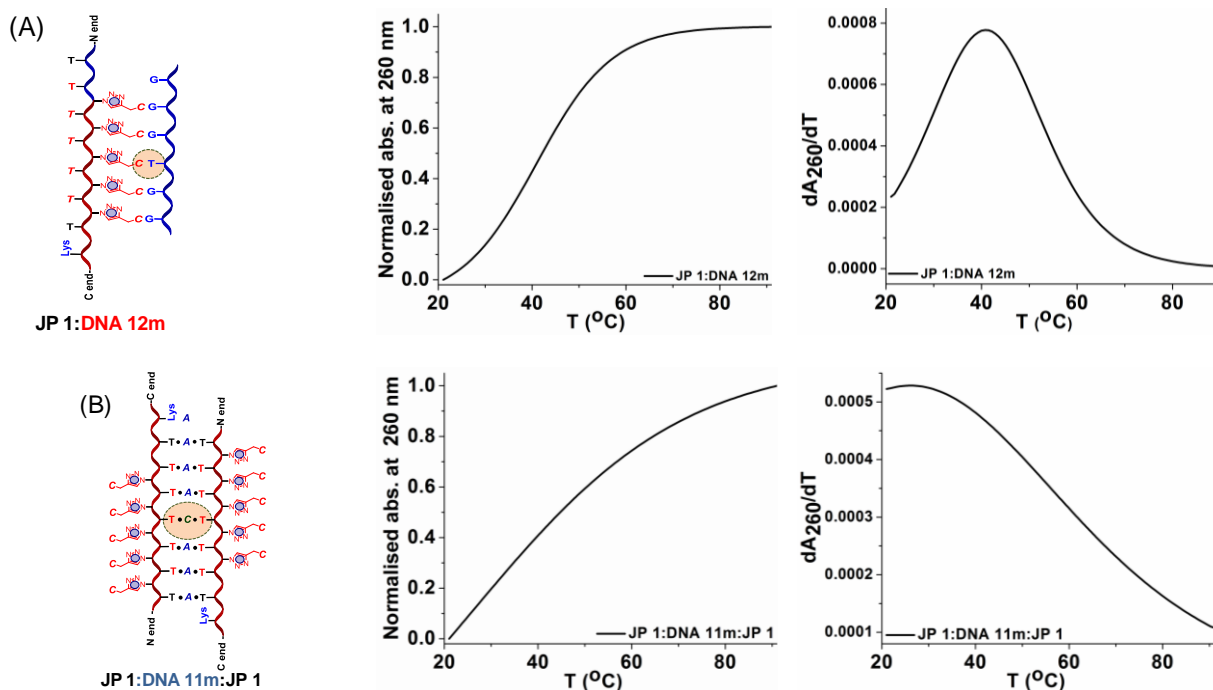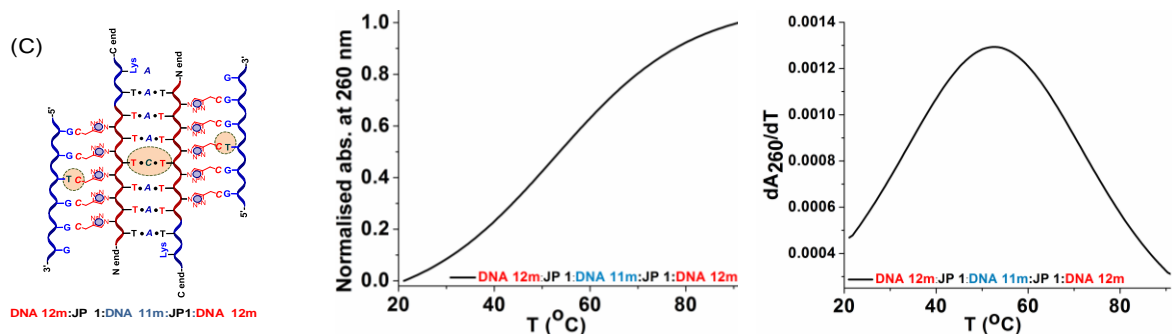
Figure 4.14 Temperature dependent UV absorbance curves for mismatch DNA complexes of T_7jp-tz-C_5 (**JP 1**) with DNA **11m** (5'-AAA**C**AAAA-3') and DNA **12m** (5'-GG**T**GGG-3'); Buffer 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM (**A**) duplex **JP 1**:DNA **12m** (**B**) triplex **JP 1**:DNA **11m**:**JP 1** (**C**) duplex of triplex DNA **12m**:**JP 1**:DNA **11m**:**JP 1**:DNA **12m**, T_m values are accurate to ± 1.0 °C.

Table 4.5 UV - T_m (°C) of **JP 1** (T_7jp-tz-C_5) with mismatch DNAs

Entry	*Homo Janus* PNA:mismatch DNA complexes	T_m (°C)		ΔT_m (°C) (double duplex - T_m duplex)		ΔT_m^1 (°C) (mismatch duplex – perfect duplex)	
		Amide	Triazole	Amide	Triazole	Amide	Triazole
1	**JP 1**:DNA **12m**	-	40.8	-	-		-9.9
2	**JP 1**:DNA **11m**:**JP 1**	28.0	-	-	-	-2.2	
3	DNA **12m**:**JP 1**:DNA **11m**: **JP1**:DNA **12m**	53.2		+25.2	+12.4		-20.3

4.4.2c CD spectra of *homo Janus* triplexes and duplexes.

The CD spectra of *Janus* PNA complex with dA_8 (C_5-tz-jpT_7:dA_8:T_7jp-tz-C_5), showed positive bands at 262 nm and 282 nm that is characteristic of the (PNA-T_8)$_2$:poly dA triplex and right-handed helix,[18] as shown in (Figure 4.15A) for unmodified PNA **15**$_2$:DNA **1** triplex. In addition, a negative minimum at 249 nm with cross-over points at 240 nm and 259 nm were seen. This CD profile suggests the formation of triplex dA_8:(T_7jp-tz-C_5)$_2$ with DNA **1**:**JP 1** (1:2) stoichiometry as seen by a single transition in UV-T_m.
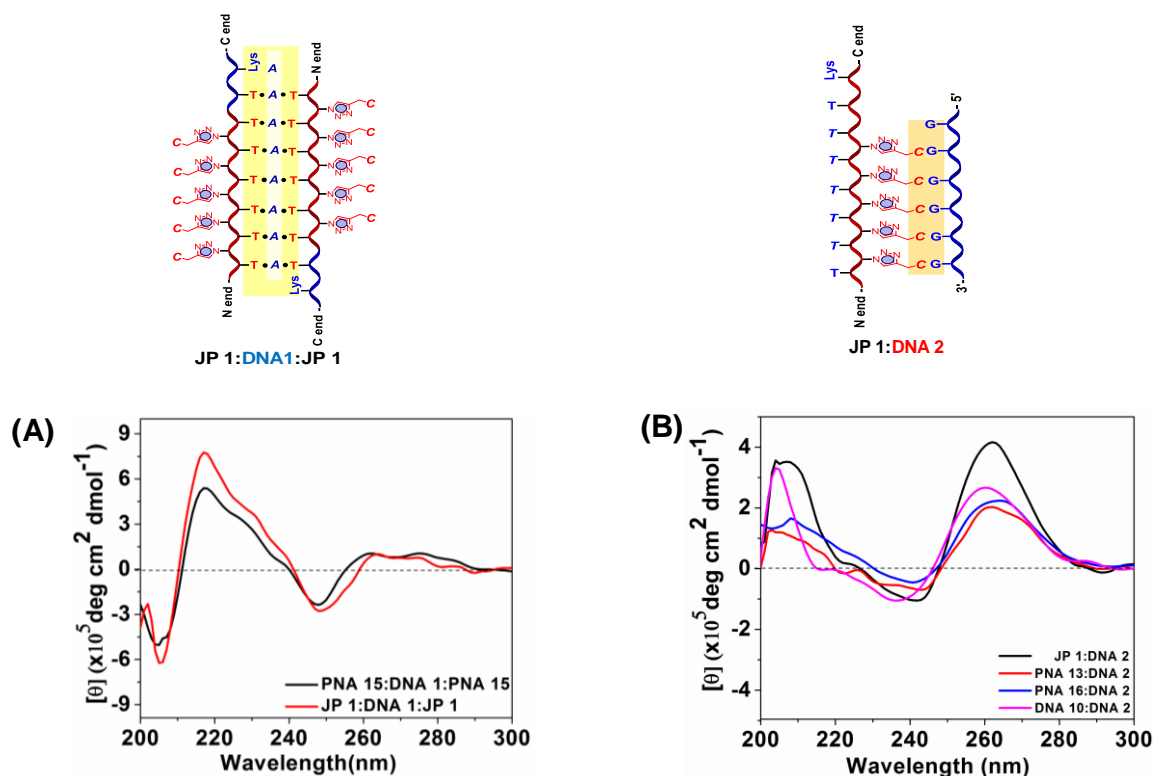
Figure 4.15 CD spectra of **(A)** Amide face DNA complexes from *aeg* PNA, *Janus* PNA T$_\gamma$*jp-tz-C$_5$* (**JP 1**) and **(B)** Triazole face DNA duplexes from PNA, *aeg-tz* PNA, *Janus*-PNA T$_\gamma$*jp-tz-C$_5$* (**JP 1**) Buffer: 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM.

CD spectra of the *Janus* PNA triazole duplex T$_\gamma$*jp-tz-C$_5$*:dG$_6$ (**JP 1**:DNA **2**) were compared with CD spectra of control duplexes *p*7-*tz-C$_5$*:dG$_6$ (*aeg-tz*-PNA **13**:DNA **2**), dC$_6$:dG$_6$ (DNA **10**:DNA **2**) and PNA-C$_5$:dG$_6$ (PNA **16**:DNA **2**) shown in (Figure 4.15B). All duplexes showed a positive band around 260 nm (± 2 nm) and a negative band around 240 nm with cross-over points at 247 nm. Further dC$_6$ (DNA **10**) / PNA-C$_5$ (PNA **16**) / *p*7-*tz-C$_5$* (PNA **13**) /T$_\gamma$*jp-tz-C$_5$* (**JP 1**) form duplexes with dG$_6$ (DNA **2**) with 1:1 stoichiometry as indicated by UV-Job plot and ITC (see later).

In comparison with the CD spectra of isolated triplexes and duplexes as shown above, the *double duplex of Janus triplex* (DNA **2**:**JP 1**:DNA **1**:**JP 1**:DNA **2**) constituted from T$_\gamma$*jp-tz-C$_5$* (**JP 1**), dA$_8$ (DNA **1**) and dG$_6$ (DNA **2**) (Figure 4.16) shows a positive maxima around 220 nm and 263 nm (shoulder at 228 nm) and a negative maxima at 245 nm with cross-over points at 237 and 253 nm. The CD spectra appears to be a composite of the amide triplex (Figure 4.15A) and triazole duplex (Figure 4.15B). In UV-T_m plot, this complex showed two T_ms with

enhanced T_m for both triplex and duplex component. Thus this CD is not a result of simple addition spectra of the isolated triplex and duplex, but arises from formation of double duplex of a triplex.
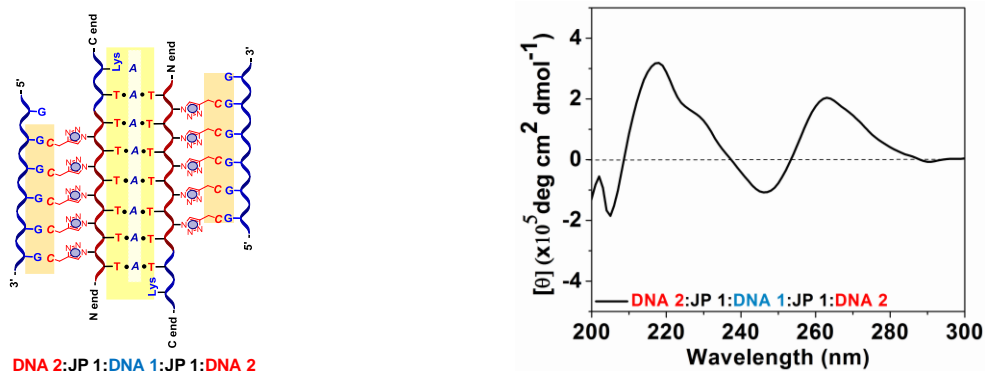
Figure 4.16 CD spectra of duplex of triplex dG$_6$:C_5-*tz*-*jp*T$_7$:dA$_8$:T$_7$*jp*-*tz*-C_5:dG$_6$ (DNA **2**:JP **1**:DNA **1**:JP **1**:DNA **2**) complex. Buffer: 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM.

CD spectral study of order of duplex formation: Whether the order of triplex / duplex formation matters to achieve the final observed CD profile of double duplex of triplex was examined by studying CD patterns in a sequential manner: (i) *homo Janus* PNA T_7jp-*tz*-C_5 (**JP 1,** 10 µM) in sodium cacodylate buffer and NaCl at pH 7.1 showed CD with a weak positive band in the region 240 nm – 260 nm and negative band centered around 270 and 285 nm. (ii) it was mixed with stoichiometric amount of the complementary dG$_6$, (DNA **2**) after which it was kept for equilibration for 10 mins and the CD spectra was recorded for T_7jp-*tz*-C_5:dG$_6$ (iii) this was followed by addition of stoichiometric amount of dA$_8$ (DNA **1**), kept for equilibrated for 10 min after each addition and the CD spectrum of dG$_6$:C_5-*tz*-*jp*T$_7$:dA$_8$:T$_7$*jp*-*tz*-C_5:dG$_6$ (DNA **2**:JP **1**:DNA **1**:JP **1**:DNA **2**) was recorded (Figure 4.17A).

The sequential CD experiment was repeated by reversing the order of DNA additions to **JP 1** first addition of dA$_8$ (DNA **1**) to generate C_5-*tz*-*jp*T$_7$:dA$_8$:T$_7$*jp*-*tz*-C_5, (**JP 1**:DNA **1**:JP **1**) triplex (followed by addition of dG$_6$ (DNA **2**) to obtain the duplex of triplex dG$_6$:C_5-*tz*-*jp*T$_7$:dA$_8$:T$_7$*jp*-*tz*-C_5:dG$_6$ (DNA **2**:JP **1**:DNA **1**:JP **1**:DNA **2**) (Figure 4.17B). It is seen that in both experiments the CD spectra of final product dG$_6$:C_5-*tz*-*jp*T$_7$:dA$_8$:T$_7$*jp*-*tz*-C_5:dG$_6$ (DNA **2**:JP **1**:DNA **1**:JP **1**:DNA **2**) blue band was identical (Figure 4.18) in terms of spectral bands and intensities. Thus, the order in which the duplexes are formed from both faces does not affect the conformational state of final double duplex of triplex.

(A) *First complex with dG$_6$ followed by dA$_8$*

T$_7$*jp-tz-C$_5$* + dG$_6$ = T$_7$*jp-tz-C$_5$*:dG$_6$; T$_7$*jp-tz-C$_5$*:dG$_6$ + dA$_8$ = dG$_6$:*C$_5$-tz-jp*T$_7$:dA$_8$:T$_7$*jp-tz-C$_5$*:dG$_6$

(B) *First complex dA$_8$ followed by dG$_6$*

T$_7$*jp-tz-C$_5$* + dA$_8$ = *C$_5$-tz-jp*T$_7$:dA$_8$:T$_7$*jp-tz-C$_5$*; *C$_5$-tz-jp*T$_7$:dA$_8$:T$_7$*jp-tz-C$_5$* + dG$_6$ = dG$_6$:*C$_5$-tz-jp*T$_7$:dA$_8$:T$_7$*jp-tz-C$_5$*:dG$_6$
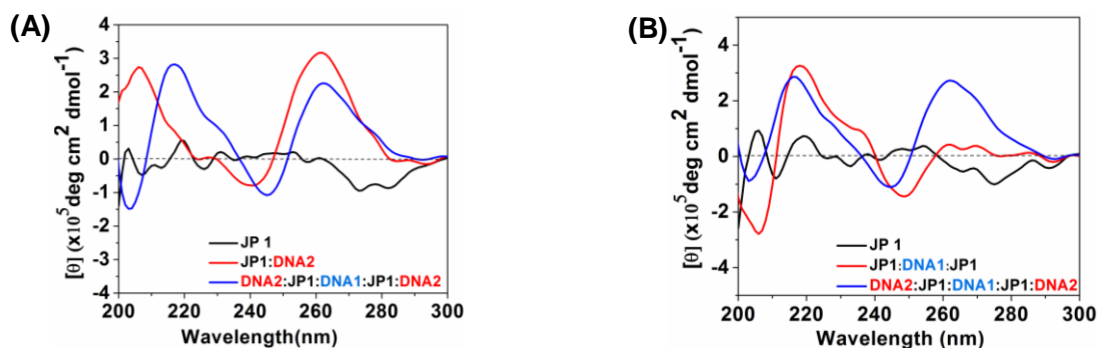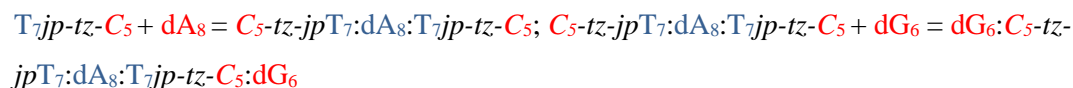
Figure 4.17 CD spectra at 10 °C **(A)** T$_7$*jp-tz-C$_5$* (**JP 1,** 10 μM); T$_7$*jp-tz-C$_5$*+ triazole face complementary dG$_6$, (DNA **2,** 10 μM) to obtain duplex T$_7$*jp-tz-C$_5$*:dG$_6$ (**JP 1**:DNA **2**) (1:1); duplex T$_7$*jp-tz-C$_5$*:dG$_6$ + amide amide face complementary dA$_8$ (DNA **1**, 10 μM) to obtain double duplex of triplex dG$_6$:*C$_5$-tz-jp*T$_7$:dA$_8$:T$_7$*jp-tz-C$_5$*:dG$_6$ (DNA **2**:**JP 1**:DNA **1**:**JP 1**:DNA **2**, 1:1:1). **(B)** T$_7$*jp-tz-C$_5$* (**JP 1,** 10 μM); T$_7$*jp-tz-C$_5$* + amide face complementary dA$_8$, (DNA **1**, 10 μM) to obtain triplex *C$_5$-tz-jp*T$_7$:dA$_8$:T$_7$*jp-tz-C$_5$* (**JP 1**:DNA **1**:**JP 1,** 1:1); triplex *C$_5$-tz-jp*T$_7$:dA$_8$:T$_7$*jp-tz-C$_5$* + triazole face complementary dG$_6$ (DNA **2**, 10 μM) to obtain double duplex of triplex dG$_6$:*C$_5$-tz-jp*T$_7$:dA$_8$:T$_7$*jp-tz-C$_5$*:dG$_6$ (DNA **2**:**JP 1**:DNA **1**:**JP 1**:DNA **2**, 1:1:1)
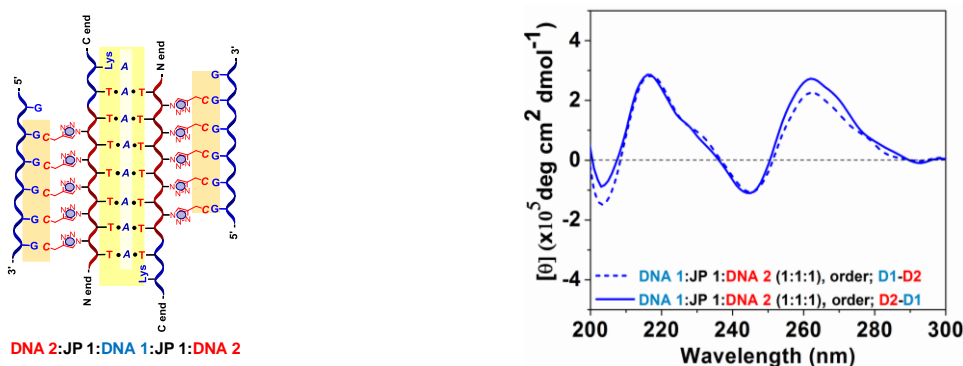
Figure 4.18 Final spectra of dG$_6$:*C$_5$-tz-jp*T$_7$:dA$_8$:T$_7$*jp-tz-C$_5$*:dG$_6$ (DNA **2**:**JP 1**:DNA **1**:**JP 1**:DNA **2**) from figure 4.17A, spectrum (blue) and 4.17B (dash blue)

4.4.2d ESI-MS of *homo Janus* PNA:DNA duplex/double duplex

Non-covalent complex formation of *Janus* PNA:DNA duplexes and triplexes was also demonstrated by using ESI-MS. The samples of *Janus* PNAs along with complementary DNAs were individually injected into ESI-MS sample chamber. The *Janus* PNA, T$_7$*jp-tz-C$_5$* (**JP 1**) was

injected along with dA_8 (DNA **1**, amide face) and dG_6 (DNA **2**, triazole face) and the mass spectra of the complex showed a peak at m/e 7807.9698 (Figure 4.19). This corresponds to a calculated mass of 7808.3419 for molecular formula $[C_{272}H_{368}N_{131}O_{123}P_{12}]$ of the complex with a composition dA_8:T_7jp-tz-C_5:dG_6 + 2xAcOH + 13 H_2O (DNA **1**:JP **1**:DNA **2**). This suggests the formation of double duplex under the gas phase electrospray ionizing conditions of mass spectrometer.
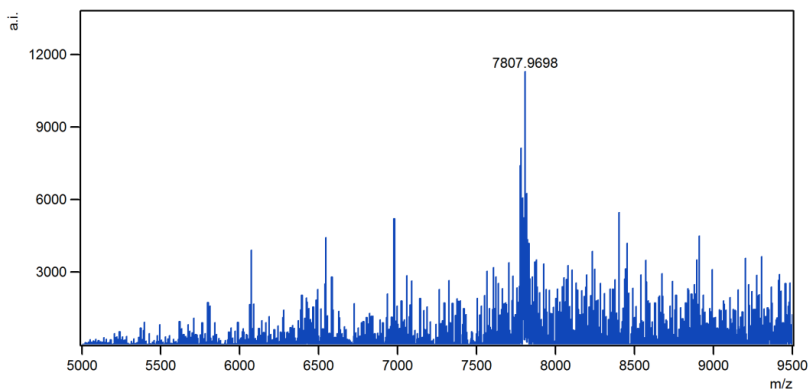
Figure 4.19 ESI-MS of (dA_8 + T_7jp-tz-C_5 + dG_6)

Similar mass spectral experiments using stoichiometric amounts of T_7jp-tz-C_5 (**JP 1**) + dA_8 (DNA **1**) indicated formation of (*Janus* PNA$_2$:DNA **1**) complex corresponding to triplex C_5-tz-jpT$_7$:dA_8:T_7jp-tz-C_5 (**JP 1**:DNA **1**:**JP 1**, Figure 4.20). Experiments with T_7jp-tz-C_5 (**JP 1**) + dG_6 (DNA **2**) gave mass spectral peak corresponding to *Janus* PNA:DNA duplex T_7jp-tz-C_5:dG_6 (Figure 4.21). Overall, ESI-MS data gave additional proof for the formation of duplex, triplex and double duplexes from *Janus* PNAs in gas phase vapor conditions. Observation of mass peaks corresponding to specific compositions supports formation of duplexes and triplexes under gas phase mass spectral conditions.
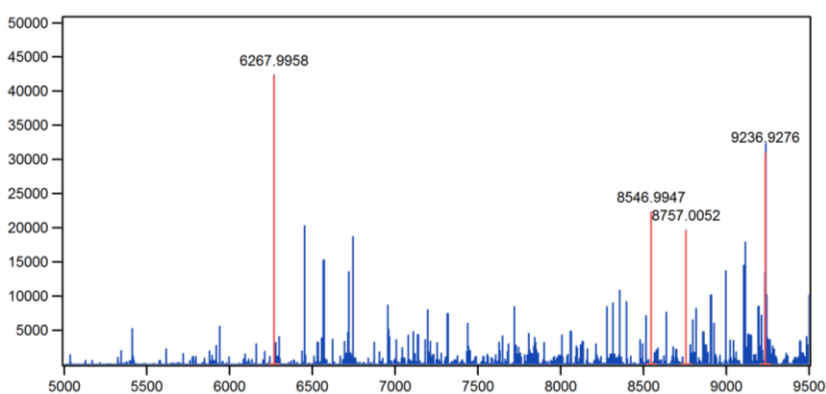
Figure 4.20 ESI-MS of T_7jp-tz-C_5 + dA_8 + T_7jp-tz-C_5 (**JP 1**:DNA **1**:**JP 1**)

ESI-MS of triplex M = [T$_7$*jp-tz-*C$_5$ + dA$_8$ + T$_7$*jp-tz-*C$_5$**]**

Obs. [M*+K]$^+$ = 9236.9276 Calc. [M*+K]$^+$ = 9236.2497

where M*= C$_5$-*tz-jp*T$_7$:dA$_8$:T$_7$*jp-tz-*C$_5$ + 9 AcOH + H$_2$O [C$_{354}$H$_{461}$N$_{162}$O$_{125}$P$_7$]

ESI-MS of duplex M = [T$_7$*jp-tz-*C$_5$ + dA$_8$**]**

Obs. [M*+H]$^+$ = 6267.9958 Calc. [M*+H]$^+$ = 6268.1399

where M*= T$_7$*jp-tz-*C$_5$:dA$_8$ + 8 AcOH + 7 H$_2$O + 3 ACN [C$_{233}$H$_{318}$N$_{101}$O$_{95}$P$_7$]
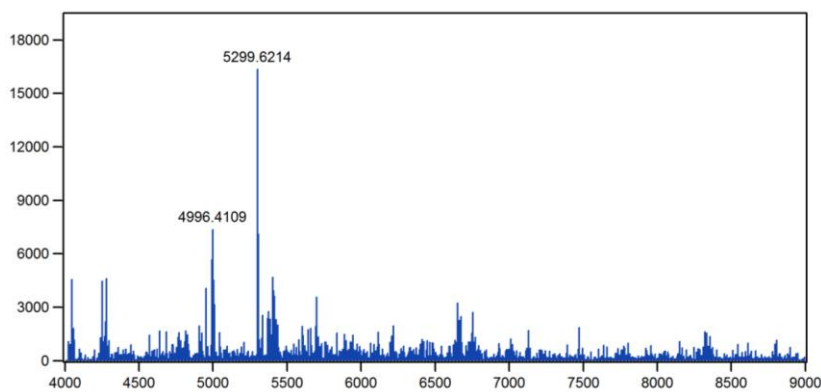
Figure 4.21 ESI-MS of T$_7$*jp-tz-*C$_5$ + dG$_6$ (**JP 1:**DNA **2**)

ESI-MS of duplex M = [T$_7$*jp-tz-*C$_5$ + dG$_6$**]**

Obs. [M* + K]$^+$ = 5299.6214, Calc. [M* + K]$^+$ 5298.8101

Where M* = T$_7$*jp-tz-*C$_5$ + dG$_6$ + 6 ACN + K [C$_{200}$H$_{254}$KN$_{97}$O$_{68}$P$_5$]

4.4.2e ITC study of complexes of *homo Janus* PNA T$_7$*jp-tz-*C$_5$ (JP 1) with complementary DNAs

This section describes the ITC binding experiments of *Homo Janus* PNA T$_7$*jp-tz-*C$_5$ (**JP 1**) to complementary dA$_8$ (DNA **1**) to amide face and with dG$_6$ (DNA **2**) complementary to triazole face, to examine the differential thermodynamics of binding of *Janus* PNAs to amide and triazole faces.

The binding isotherm of PNA T$_7$*jp-tz-*C$_5$ (**JP 1**, 81.0 μM) with dA$_8$ (DNA **1,** 1.05 mM) showed a sigmoidal profile (Figure 4.22) suggesting co-operative binding. The stoichiometry of binding N of 0.5, (N = Ratio of No. of nucleobases involved in binding from DNA / No. of nucleobases involved in binding from Janus PNA, 7/14 = 0.5)**,** suggested formation of triplex C$_5$-*jp-tz*T$_7$:dA$_8$:T$_7$*jp-tz-*C$_5$ (**JP 1**$_2$:DNA **1**) with enthalpy ΔH of - 28.4 kcal/mol and K$_D$ of 0.34 x

10^{-6} M. Thus ITC results supported the observation of *Janus* PNA$_2$:DNA triplex as indicated by CD spectroscopy and ESI-MS.
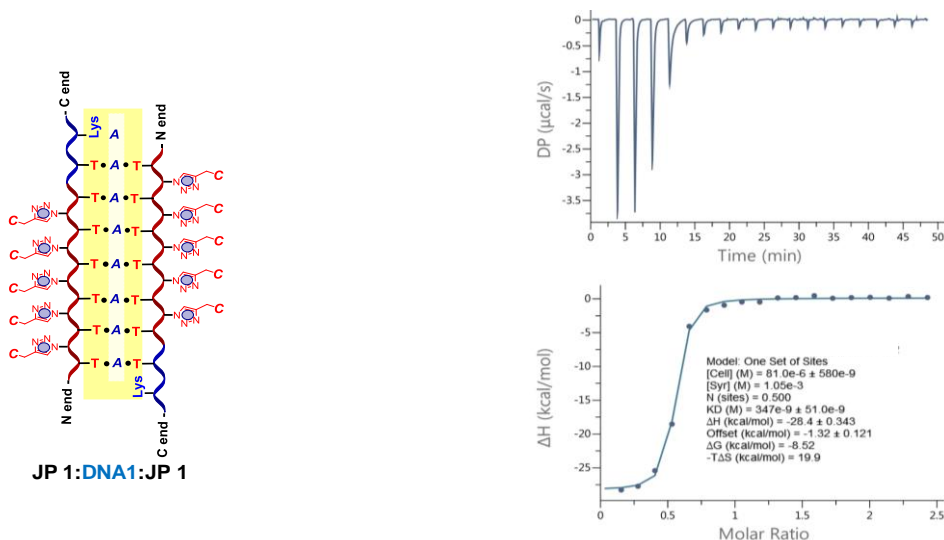
Figure 4.22 ITC data for **JP 1**$_2$:DNA **1** triplex formation

In a second experiment, binding of T$_7$*jp-tz*-C$_5$ (**JP 1,** 81.0 μM) with dG$_6$ (DNA **2,** 0.75 mM) that is complementary to triazole face also showed a sigmoidal (Figure 4.23) behavior. The thermodynamic parameters extracted from the titration data corresponded to ΔH of -28.4 kcal/mol; K$_D$, 2.0 x 10^{-6} M and N of 1.08, (N = Ratio of No. of nucleobases involved in binding from DNA / No. of nucleobases involved in binding from Janus PNA, 10/10 = 1.0), fitting to formation of **JP 1**:DNA **2** (1:1) duplex. The negative ΔG (-8.52 kcal/mol) indicated an overall favorable association.
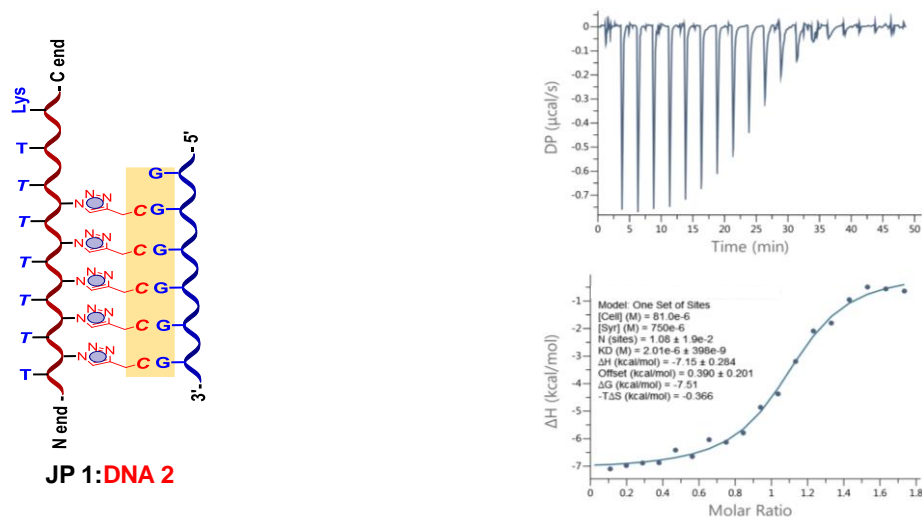
Figure 4.23 ITC data for binding of **JP 1**:DNA **2** duplex formation

The formation of the *Janus* double duplex of triplex was examined by titrating the Janus PNA T$_7$*jp-tz-C$_5$* (**JP 1**) with mixture of the complementary dA$_8$ (DNA **1**) and dG$_6$ (DNA **2**). Since ITC experiment cannot be done with sequential titration with two binding components, converting a trimolecular reaction, in principle into bimolecular equivalent is possible by using the mixture of cDNAs as titre to generate duplex of triplex *dG$_6$:C$_5$-tz-jp*T$_7$*:dA$_8$:*T$_7$*jp-tz-C$_5$:dG$_6$* (DNA **2**:**JP 1**:DNA **1**:**JP 1**:DNA **2**).
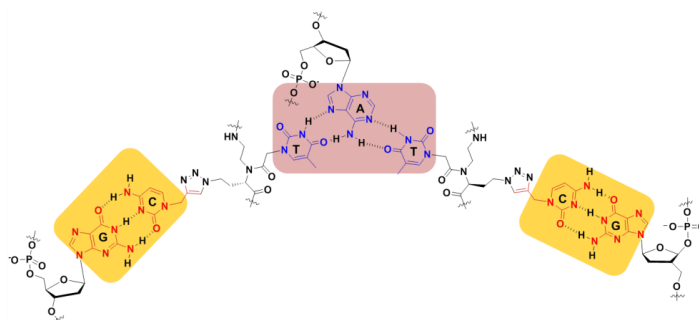
Figure 4.24 Representation of double duplex of triplex *dG$_6$:C$_5$-tz-jp*T$_7$:dA$_8$:T$_7$*jp-tz-C$_5$:dG$_6$* (DNA **2**:DNA **1**:**JP 1**:DNA **2**)

When a mixture of dA$_8$ (DNA **1,** 1.05 mM of 7 nucleobases in DNA, effective concn/nucleobase is 150 µM) and dG$_6$ (DNA **2,** 1.5 mM of 5 nucleobases in DNA, effective concn/nucleobase is 300 µM) (DNA **1**:DNA **2** = 1:2) was titrated with Janus PNA **JP 1** (81.0 µM), a second transition was distinctly observed near molar ratio of 2. The ITC data is best fitted to two-set of sites (transition) model (Figure 4.25) showing very similar ITC curve as reported[28]. The magnitude of ΔH_1 (-10.6 kcal/mol) was less compared to that observed for the individual triplex (**JP 1**)$_2$:DNA **1** (-28.4 kcal/mol) and slightly more than that for duplex **JP 1**:DNA **2** (-7.15 kcal/mol). ΔH_2 (-2.78 kcal/mol) for the duplex component decreased compared to corresponding individual duplex. It is clearly seen the K_{D1} (0.2 ± 0.07 x 10^{-6} M) decreased compared K_D (0.35 x 10^{-6} M) of (**JP 1**)$_2$:DNA **1** amide face triplex. However the K_{D2} (2.51 ± 0.8 x 10^{-6} M) of complex DNA **1**:**JP 1**:DNA **2** was not much different from K_D (2.0 x 10^{-6} M) of individual **JP 1**:DNA **2** triazole face duplex. These results suggest the formation of a stable complex presence of both face complementary DNAs, leading to double duplex of triplex. Various experimental data are summarized in Table 4.6.
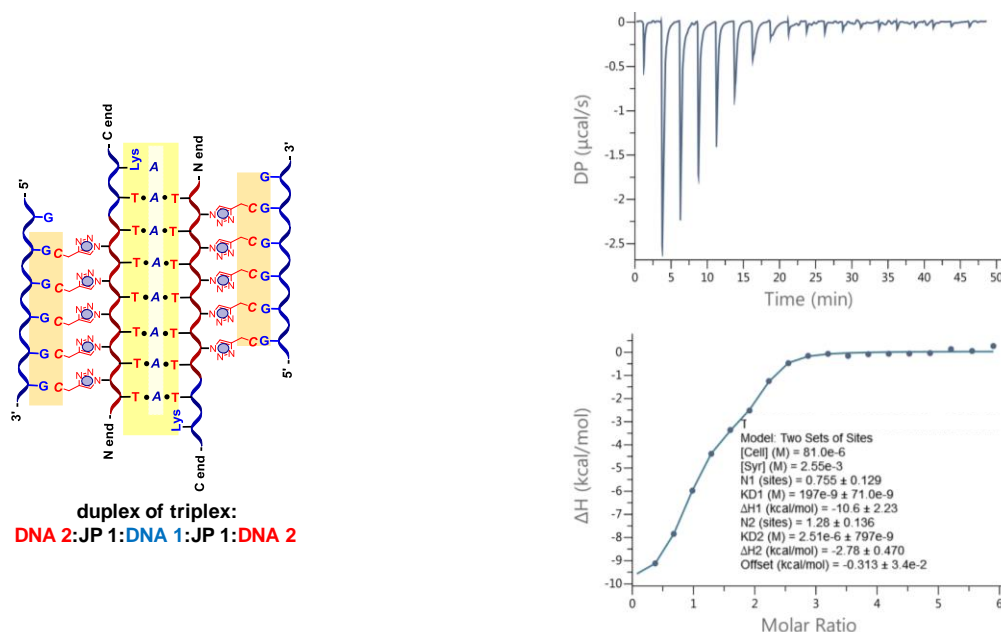
Figure 4.25 ITC data for binding of **JP 1** to mix of [DNA **1** + DNA **2** (1:2) ratio]

Table 4.6 Summary of ITC binding of *homo Janus* PNA, T$_7$*jp-tz-C$_5$* (**JP 1**) to complementary **DNA** complex

Entry	*Janus* PNA:DNA	ΔG kcal/mol		ΔH kcal/mol		ΔS cal/mol/K		K$_D$ M (x10^{-6})		N
		Amide	**Triazole**	**Amide**	**Triazole**	**Amide**	**Triazole**	**Amide**	**Triazole**	
1	**JP 1:DNA 1:JP 1** (triplex)	-8.52 ± 0.17	-	-28.4 ± 0.34	-	-69.0 ± 0.59	-	0.35 ± 0.05	-	0.5
2	**JP 1:DNA 2** (duplex)	-	-7.51 ± 0.23	-	-7.15 ± 0.28	-	+1.2 ± 0.8	-	2.0 ± 0.4	1.08
	Two set of sites model: mixture of dA$_8$ + dG$_6$ (1:2)									
3*	DNA 2:JP 1:DNA 1:JP 1:DNA 2 (Double duplex of triplex)	-8.85 ± 0.43	-	-10.6 ± 2.23		-6.1 ± 1.5	-	0.2 ± 0.07	-	0.76
		-	-7.39 ± 0.38	-	-2.78 ± 0.47	-	+15.9 ± 1.3	-	2.51 ± 0.8	1.28

N = No. of nucleobase involved in binding from DNA / No. of nucleobase involved in binding from *Janus* PNA

4.4.3 Binding studies of *Chimeric Janus* PNA m*jp-tz-C$_5$* with complementary DNAs

The thermal stability of duplex and double duplexes from *chimeric Janus* PNA m*jp-tz-C$_5$* (having hetero nucleobases on amide face and homo cytosine on triazole face) with complementary DNA and mismatched DNA was determined by temperature dependent UV

absorbance. Conformational studies were done by CD spectroscopy and further proof for formation of double duplex was provided by ESI-MS.

4.4.3a Thermal stability of *chimeric Janus* PNA JP 2 (m*jp-tz-C₅*):cDNA duplexes

The *chimeric Janus* PNA m*jp-tz-C₅* (**JP 2**) was hybridized individually with two different complementary DNAs, dG₆ (DNA **2** for triazole face) and DNA **3** (5'-GTGATCT-3') for the amide face mixed sequence shown in (Figure 4.26). The triazole face duplex m*jp-tz-C₅*:dG₆ (**JP 2**:DNA **2**) showed a single transition corresponding to T_m = 51.2 °C (Table 4.7, entry 1), similar to T_m of 50.7 °C observed previously for *homo Janu*s PNA duplex T₇*jp-tz-C₅*:dG₆ (**JP 1**:DNA **2**). The duplex from amide face *mixed sequence* cDNA **3**:m*jp-tz-C₅* (DNA **3**:**JP 2**) showed a single transition with T_m of 33.6 °C (Table 4.7, entry 2). In simultaneous presence of complementary DNA **2** and DNA **3**, the *chimeric Janus* exhibited two melting transitions with T_{m1} = 38.8 °C and T_{m2} = 77.1 °C PNA indicating formation of double duplex cDNA **3**:m*jp-tz-C₅*:dG₆ (DNA **3**:**JP 2**:DNA **2**) (Table 4.7, entry 3). The pattern was similar to that observed previously for double duplex of triplex dG₆:*C₅-tz-jp*T₇:dA₈:T₇*jp-tz-C₅*:dG₆ (DNA **2**:**JP 1**:DNA **1**:**JP 1**:DNA **2**). Compared with their individual isolated duplexes, the T_m of the amide face duplex in double duplex was enhanced by +5.2 °C while that from triazole face was increased by +25.2 °C (Table 4.6). This suggested synergistic stabilizing effects exerted by each face duplex on the other duplex. *Among the duplexes from each side in both homo (**JP 1**) and chimeric (**JP 2**) Janus PNAs, the stability of the triazole face duplex was always higher than that of the amide face duplex.*
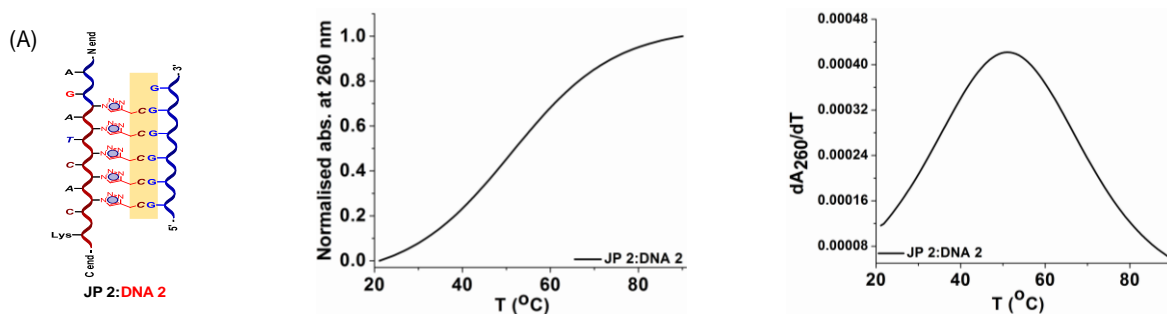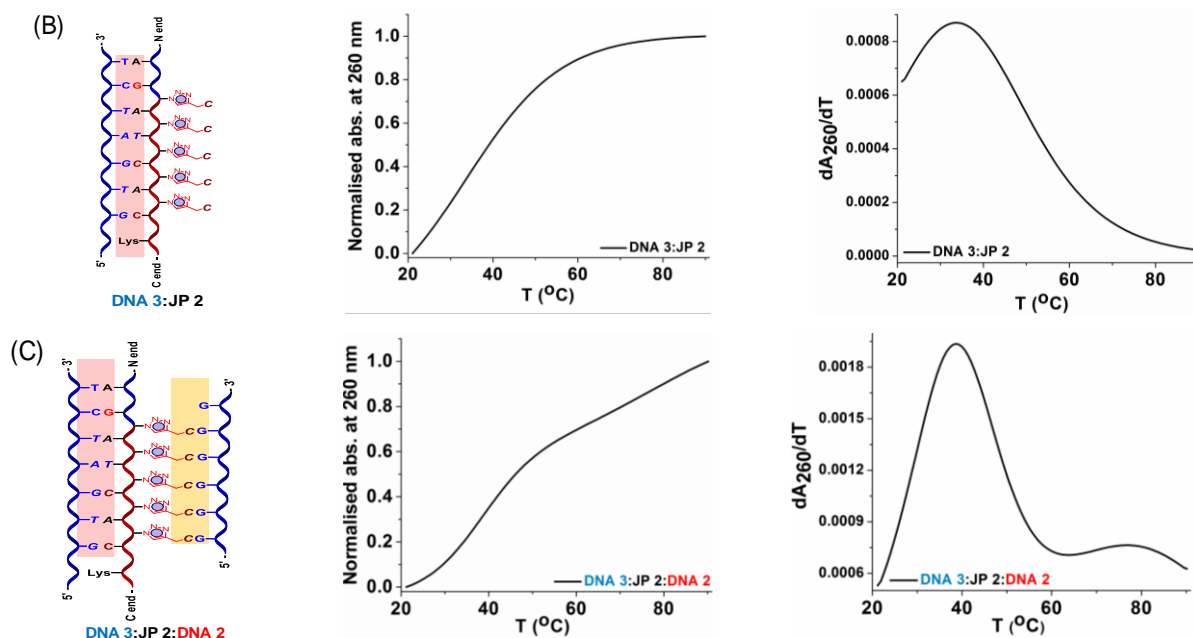
Figure 4.26 Temperature dependent UV absorbance curves for mjp-*tz*-C_5 (**JP 2**) with antiparallel complementary DNA duplex and double duplexes, DNA **2** (dG$_6$ = 5' GGGGGG 3', DNA **3** (5'-GTGATCT-3'); Buffer: 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM) **(A)** duplex mjp-*tz*-C_5:dG$_6$ (**JP 2**:DNA **2**) **(B)** duplex DNA **3**:mjp-*tz*-C_5 and **(C)** double duplex DNA **3**:mjp-*tz*-C_5:dG$_6$, T_m values are accurate to ±1.0 °C.

Table 4.7 UV-T_m (°C) of mjp-*tz*-C_5 (**JP 2**) with complementary DNA **2** and DNA **3**

Entry	*Chimeric* **PNA:DNA complexes**	T_m (°C)		ΔT_m (°C) (double duplex - T_m duplex)	
		Amide	Triazole	Amide	Triazole
1	mjp-*tz*-C_5:dG$_6$ (**JP 2:DNA 2**)	-	51.2	-	-
2	cDNA:mjp-*tz*-C_5 (DNA **3:JP 2**)	33.6	-	-	-
3	cDNA:mjp-*tz*-C_5:dG$_6$ (DNA **3:JP 2:DNA 2**)	38.8	77.1	+5.2	+25.9

Melting studies on mismatch duplexes. In order to understand the enhanced stability of each face duplex and attribute it to specific H-bonding as per WC base pairing, melting studies were done with mjp-*tz*-C_5 (**JP 2**) using DNA **14m** (dG$_2$T$_2$G$_2$) that has two C-T mismatches in triazole face duplex and DNA **13m** that has two mismatches (G-T and A-C) on amide face are shown in (Figure 4.27). The triazole face duplex derived from mismatch DNA **14m** with mjp-*tz*-C_5 (**JP 2:DNA 14m**) showed a single transition corresponding T_m = 29.3 °C (Table 4.8, entry 1), in comparison with T_m of 51.2 °C for the perfect chimeric duplex mjp-*tz*-C_5:dG$_6$ (**JP 2:dG$_6$**). Thus

(2xC-T) mismatches introduced destabilization of duplex by 21.9 ^{o}C. Similarly, the duplex from amide face DNA **13m**:m*jp-tz*-**C$_5$** having two mismatches (G-T and A-C) also showed a single transition with T_m of 27.7 ^{o}C (Table 4.8, entry 2), which is lower by 5.9 ^{o}C than the perfect duplex DNA **3:JP 2**.

In the presence of both side mismatch DNAs (DNA **13m** and DNA **14m**), the *chimeric Janus* PNA m*jp-tz*-**C$_5$** (**JP 2**) exhibited a single melting transition with T_m = 37.5 °C (Figure 4.26). Compared with their individual isolated mismatched duplexes DNA **13m:JP 2** and **JP 2**:DNA **14m**, the T_ms of both amide and triazole duplexes DNA in the ternary complex **13m:JP 2: DNA 14m** were enhanced by +9.8 ^{o}C and +8.2 ^{o}C respectively (Table 4.8, entry 3). The observance of single transition in this double duplex may be interpreted as due to degenerate T_ms of both amide face and triazole face duplexes. Such enhancement, similar to that seen in perfect duplexes substantiate the formation of double duplexes in both perfect and mismatched duplexes of *Janus* PNA by complementary base pairing.
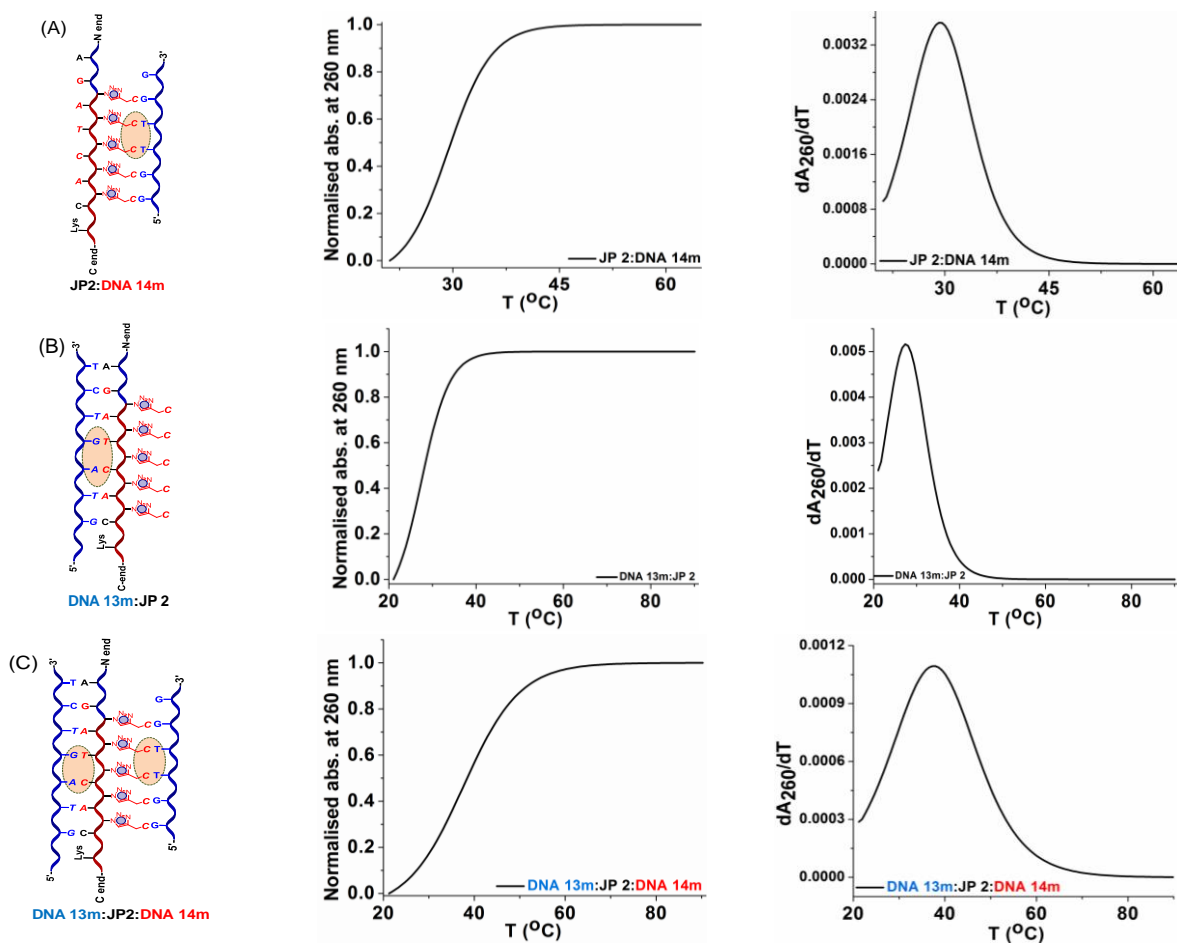
Figure 4.27 Temperature dependent UV absorbance curves for **JP 2** or m*jp-tz-C₅* with antiparallel mismatch DNA duplex and double duplexes, DNA **14m** = 5' GG**TT**GG 3', DNA **13m** = 5' GT**AG**TCT 3'; Buffer: 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM) **(A) JP 2**:DNA **14m (B)** DNA **13m**:**JP 2** **(C)** DNA **13m**:**JP 2**:DNA **14m**, T_m values are accurate to ±1.0 °C.

Table 4.8 UV - T_m (°C) of **JP 2** (m*jp-tz-C₅*) with mismatch DNA **13m** and DNA **14m**

Entry	*Chimeric Janus* PNA:mismatchDNA complexes	T_m (°C)		ΔT_m (°C) (double duplex - T_m duplex)		ΔT_m^1 (°C) (mismatch duplex – perfect duplex)	
		Amide	Triazole	Amide	Triazole	Amide	Triazole
1	**JP 2**:DNA **14m**	-	29.3	-	-		-21.9
2	DNA **13m**:**JP 2**	27.7	-	-	-	-5.9	
3	DNA **13m**:**JP 2**:DNA **14m**	37.5		+9.8	+8.2	-1.3	-39.6

ΔT_m^1 indicates the difference in T_m of m*jp-tz-C₅* with mismatch DNA and cDNA.

4.4.3b CD spectra of *chimeric Janus* PNA m*jp-tz-C₅*:DNA complexes

The CD spectra of the *chimeric Janus* PNA m*jp-tz-C₅* **JP 2** with amide face complementary DNA **2** and triazole face complementary DNA **3** to form respective duplexes are shown in (Figure 4.28). The amide face duplex DNA **3**:m*jp-tz-C₅* (DNA **3**:**JP 2**) shows a positive maxima at 282 nm and a negative maxima at 258 nm with cross-over points at 230 and 264 nm and the triazole face duplex m*jp-tz-C₅*:DNA **2** (**JP 2**:DNA **2**) shows a positive maxima at 261 nm and a negative maxima at 238 nm with cross-over points at 227 and 249 nm, similar to triazole face duplex of **JP 1.**
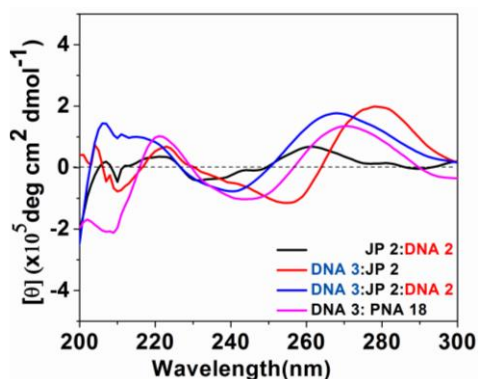
Figure 4.28 CD spectra of triazole face duplex m*jp-tz-C₅*:dG₆ (**JP 2**:DNA**2**), amide face duplex cDNA:m*jp-tz-C₅* (DNA **3**:**JP 2**), duplex of duplex cDNA:m*jp-tz-C₅*:dG₆ (DNA **3**:**JP 2**:DNA **2**), and *aeg*PNA DNA duplex, Buffer: 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM

The CD spectra of *chimeric Janus* double duplex cDNA:m*jp-tz-C₅*:dG₆ (DNA **3**:**JP 2**:DNA **2**) shows a positive maxima at 270 nm and a negative band with maxima at 241 nm with cross-over points at 225 and 250 nm (Figure 4.29). Comparison with single duplexes from amide and triazole faces, this suggests that positive band at 270 nm seems to be characteristic of double duplex as this is absent in single duplexes. For assessment, the CD spectrum of complex of *homo Janus* PNA **JP 1** with both amide and triazole faces dG₆:*C₅-tz-jp*T₇:dA₈:T₇*jp-tz-C₅*:dG₆ is also shown in Figure 4.29. In comparison, the double duplex cDNA:m*jp-tz-C₅*:dG₆ (DNA **3**:**JP 2**:DNA **2**) has positive maxima at 270 nm and a negative band with maxima at 240 nm. *Thus CD can distinguish between double duplex and double duplex of triplex.*
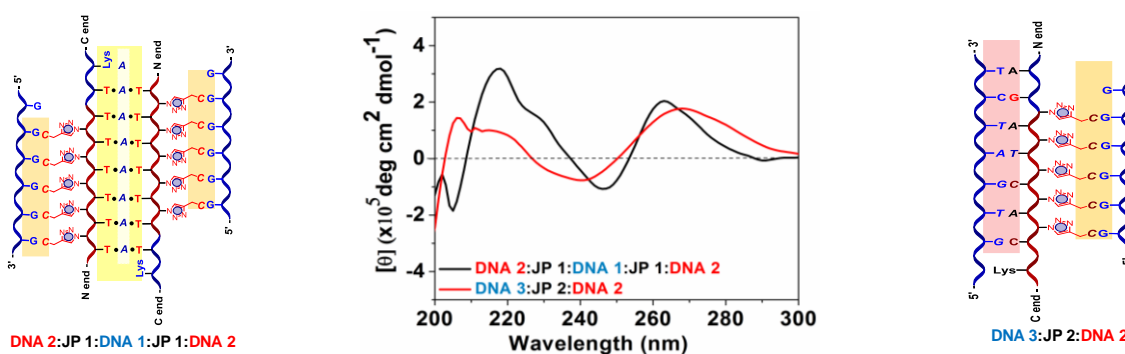
Figure 4.29 CD spectra of duplex of triplex dG₆:*C₅-tz-jp*T₇:dA₈:T₇*jp-tz-C₅*:dG₆ (DNA **1**:**JP 1**:DNA **2**:**JP 1**: DNA **1**) vs double duplex cDNA:m*jp-tz-C₅*:dG₆ (DNA **3**:**JP 2**:DNA **2**) complex. Buffer: 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM

Table 4.9 Summary of CD band pattern of T₇*jp-tz-C₅* and m*jp-tz-C₅* complex

Entry	T₇*jp-tz-C₅* (**JP 1**) complex	+ ve band (nm)	- ve band (nm)	m*jp-tz-C₅* (**JP 2**) complex complex	+ ve band (nm)	- ve band (nm)
1	*C₅-tz-jp*T₇:dA₈:T₇*jp-tz-C₅*	262 & 282	249	DNA **3**:m*jp-tz-C₅*	282	258
2	T₇*jp-tz-C₅*:dG₆	260	240	m*jp-tz-C₅*:dG₆	261	238
3	dG₆:*C₅-tz-jp*T₇:dA₈:T₇*jp-tz-C₅*:dG₆	263	245	DNA **3**:m*jp-tz-C₅*:dG₆	270	241

4.4.3c ESI-MS of *chimeric Janus* PNA:DNA double duplex

The formation of *Janus* PNA complexes in gas phase were examined by ESI-MS. The mixture of cDNA **3**+m*jp-tz-C₅*+dG₆ (DNA **3**+**JP 2**+DNA **2**) was injected into ESI-MS instrument and the mass observed (major peak) was at 8405.50 (Figure 4.30). This represents molecular formula of [$C_{271}H_{439}N_{130}O_{158}P_{11}Na$] (calculated mass is 8405.73) and corresponds to

the molecular composition of cDNA **3**:m*jp-tz*-**C₅**:**dG₆** associated with 6xAcOH+46xH$_2$O+ACN. In addition, a significant peak was seen at 7986.48 which corresponds to molecular formula C$_{269}$H$_{394}$ N$_{129}$O$_{137}$P$_{11}$ Calc. [M*+Na]$^+$ = 7968.27. The represents the complex cDNA **3**:m*jp-tz*-**C₅**:**dG₆** associated with 6xAcOH and 25xH$_2$O. The observance of these peaks *supports the formation of Janus PNA:DNA ternary double duplex cDNA 3:mjp-tz-C₅:dG₆ in agreement with transitions seen by UV-T$_m$ experiments.*
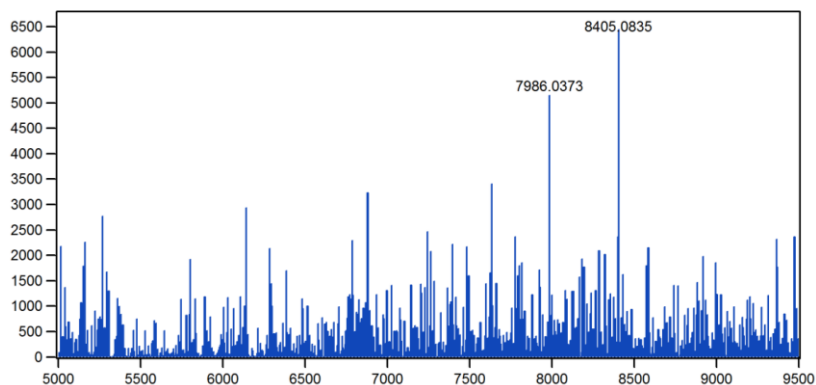
Figure 4.30 ESI-MS of cDNA **3** + m*jp-tz*-**C₅** + **dG₆** (DNA **3**:**JP 2**:DNA **2**)

4.4.4 Binding studies of *Hetero Janus* PNAs m*jp-tz₄*-*CACG* and m*jp-tz₅*-*CCACG* with complementary DNA

The fully hetero Janus PNAs m*jp-tz₄*-*CACG* (**JP 5**) and m*jp-tz₅*-*CCACG* (**JP 6**) with mixed nucleobase sequence on both amide and triazole faces were studied for formation of individual duplexes from each side with complementary DNA **3** and DNA **4** and then double duplex. This was done through variable temperature absorbance (T_m), conformation characterization by CD spectroscopy and evaluation of thermodynamic parameters by ITC experiments.

4.4.4a Thermal stability of duplexes and double duplex from *hetero Janus* PNAs m*jp-tz₄*-*CACG* (JP 5) and m*jp-tz₅*-*CCACG* (JP 6) with complementary DNAs

The *hetero Janus* PNA:DNA duplex from triazole face constituted from m*jp-tz₄*-*CACG* and DNA **4** (triazole duplex, **JP 5**:DNA **4**) showed single transition corresponding to T_m = 48.7 °C (Figure 4.31, Table 4.10, entry 1) and duplex from the amide face with DNA **3** and m*jp-tz₄*-*CACG* (amide face duplex DNA **3**:**JP 5**) gave a T_m of 30.2 °C (Table 4.10, entry 2). As with previous examples, the triazole duplex had a higher T_m (ΔT_m +18.5 °C) compared to amide

duplex. However, the double duplex DNA **3**:m*jp-tz₄-CACG*:DNA **4** (DNA **3**:**JP 5**:DNA **4**) displayed a single broad transition with a T_m of 56.6 °C (Table 4.10, entry 3), which is higher than that of individual duplexes. This indicates an overlap of transitions from duplexes of amide and triazole faces, stabilized by +26.4 °C (amide) and +7.9 °C (triazole) respectively. A shorter (tetrameric) sequence on triazole face, perhaps lead to lower T_m enhancement compared to amide face, resulting in non-resolution of two transitions.
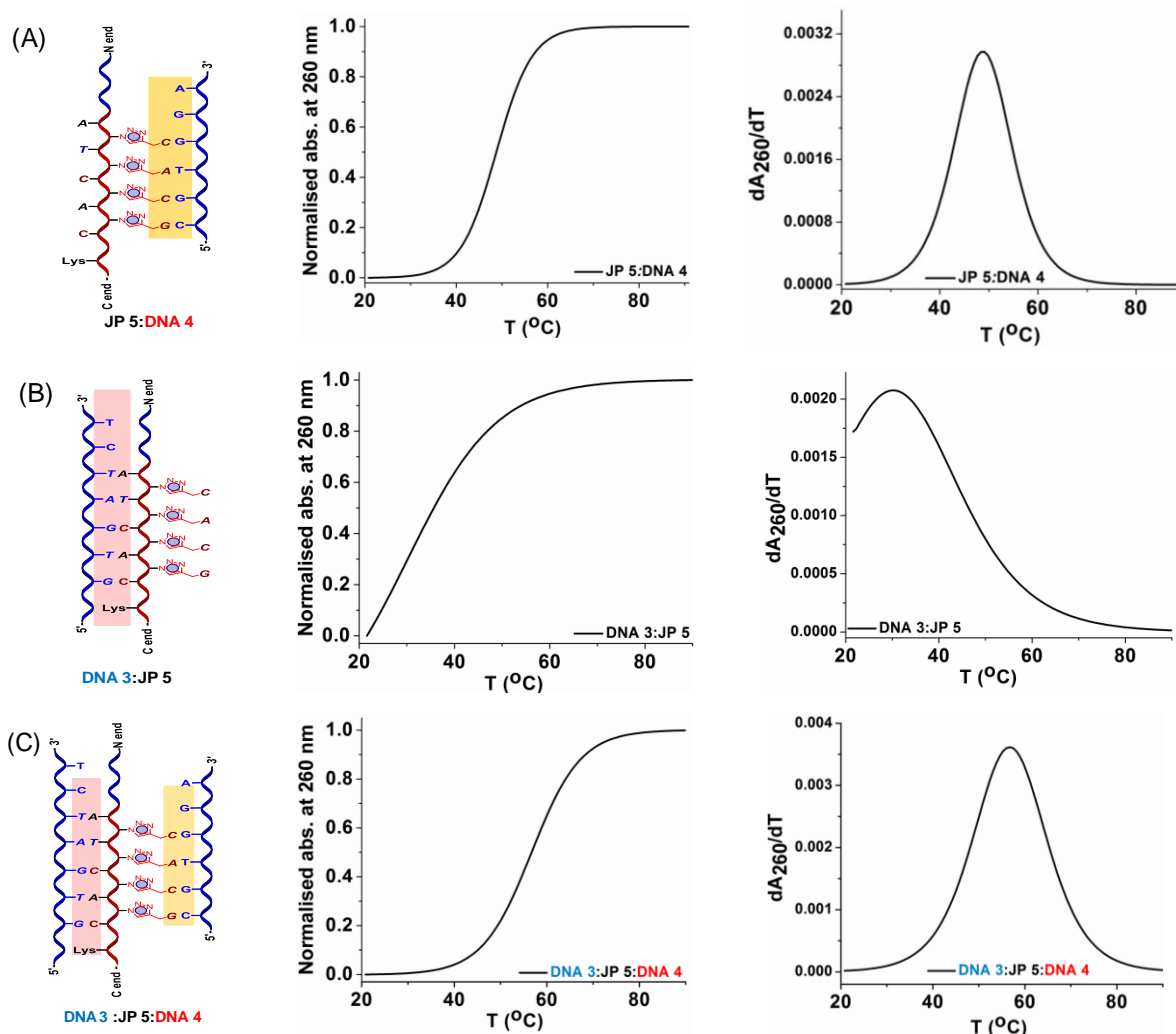
Figure 4.31 Temperature dependent UV absorbance curves for DNA duplexes from *hetero Janus* PNA m*jp-tz₄-CACG* (**JP 5**), 5'-GTGATCT-3' (DNA **3**), 5'-CGTGGA-3' (DNA **4**). Buffer: 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM) **(A)** *mjp-tz₄*-**CACG**:**DNA 4** (**JP 5**:**DNA 4**) **(B)** DNA **3**:*mjp-tz₄*-**CACG** (DNA **3**:**JP 5**) **(C)** DNA **3**:*mjp-tz₄*-**CACG**:DNA **4** (DNA **3**:**JP 5**:DNA **4**), T_m values are accurate to ±1.0 °C.

Table 4.10 UV-T_m (°C) of m*jp-tz₄*-*CACG* (**JP 5**) with complementary DNA **3** and DNA **4**

Entry	*Hetero Janus* **PNA:DNA complexes**	T_m (°C)		ΔT_m (°C) (double duplex - T_m duplex)	
		Amide	Triazole	Amide	Triazole
1	*mjp-tz₄*-CACG:DNA **4** (**JP 5**:DNA **4**)	-	48.7	-	-
2	DNA **3**:*mjp-tz₄*-CACG (DNA **3**:**JP 5**)	30.2	-	-	-
3	DNA **3**:*mjp-tz₄*-CACG:DNA **4** (DNA **3**:**JP 5**:DNA **4**)	56.6		+26.4	+7.9

To test this reasoning *hetero Janus* PNA m*jp-tz₅*-*CCACG* (**JP 6**) having a pentameric sequence on triazole face with an additional cytidine (C) was examined for duplex formation with its complementary DNA **4**. The T_m of pentameric triazole duplex m*jp-tz₅*-*CCACG*:DNA (**JP 6**:DNA **4**) now increased to 51.5 °C (Figure 4.32, Table 4.11, entry 1) while that of the amide duplex DNA **6**:m*jp-tz₅*-*CCACG* (DNA **6**:**JP 6**) was 32.3 °C (Table 4.11, entry 2). The *hetero Janus* PNA double duplex DNA **6**:m*jp-tz₅*-*CCACG*:DNA **4** now exhibited two transitions: the amide duplex with almost similar stability (32.3 °C as expected) and the triazole duplex with a T_m of 79.3 °C, stabilized significantly by +28.8 °C over T_m of single duplex (Table 4.11, entry 3). This confirmed the validity of previous reasoning on observance of a single transition in Janus PNA tetrameric tetrazole complex with DNA and increasing the length lead to resolution of second transition due to a higher T_m. This clearly supported premise that the relative stabilization of two duplexes in *Janus* double duplex depends on the length and nature of sequences on triazole face.
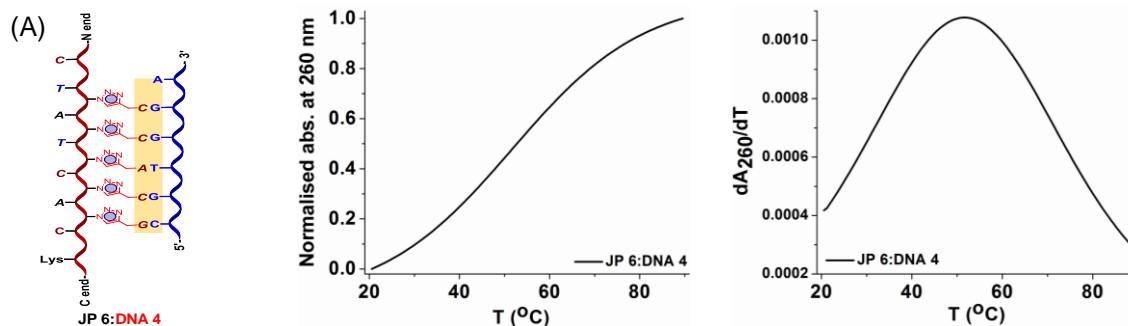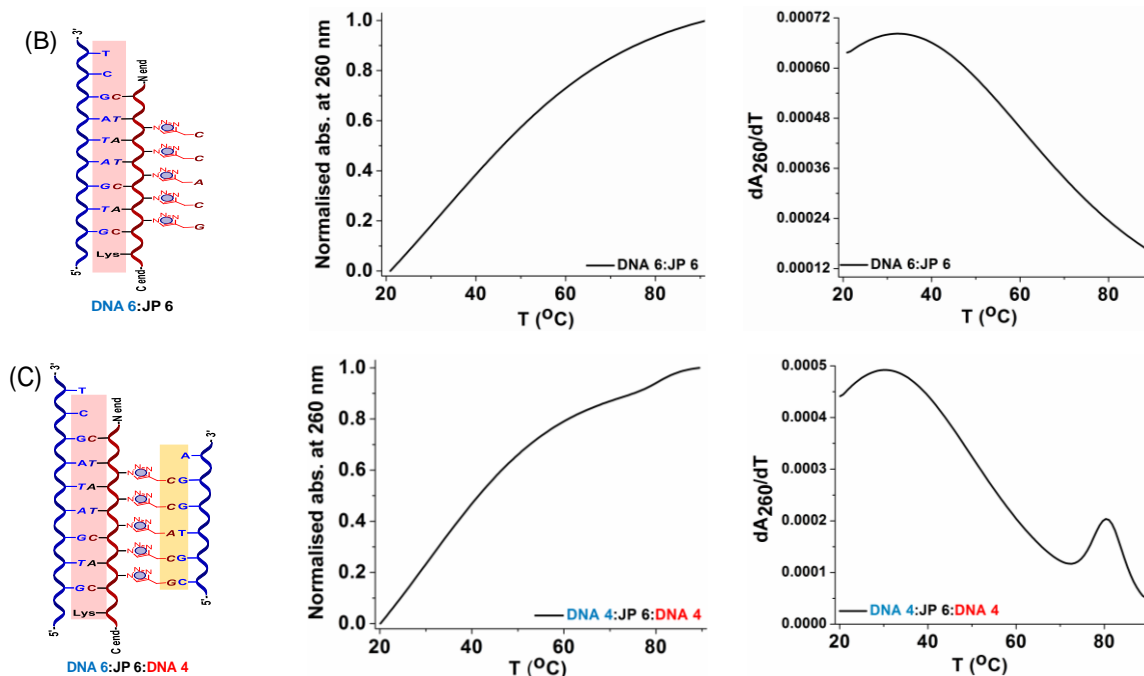
Figure 4.32 Temperature dependent UV absorbance curves for DNA duplexes of *hetero Janus* PNA m*jp-tz₅-CCACG* (**JP 6**) with DNA **4** (5'-CGTGGA-3'), DNA **6** (5'-GTGATAGCT-3'), Buffer: 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM **(A)** m*jp-tz₅-CCACG*:DNA **4** (**JP 6**:**DNA 4**) **(B)** DNA **6**:m*jp-tz₅-CCACG* (DNA **6**:**JP 6**) **(C)** DNA **6**:m*jp-tz₅-CCACG*:DNA **4** (DNA **6**:**JP 6**:**DNA 4**), T_m values are accurate to ±1.0 °C.

Table 4.11 UV-T_m (°C) of duplexes of m*jp-tz₅-CCACG* (**JP 6**) with DNA **4** and DNA **6**

Entry	*Hetero Janus* PNA:DNA complexes	T_m (°C)		ΔT_m (°C) (double duplex - T_m duplex)	
		Amide	Triazole	Amide	Triazole
1	m*jp-tz₅-CCACG*:DNA **4** (**JP 6**:DNA **4**)	-	51.5	-	-
2	DNA **6**:m*jp-tz₅-CCACG* (DNA **6**:**JP 6**)	32.3	-	-	-
3	DNA **6**:m*jp-tz₅-CCACG*:DNA **4** (DNA **6**:**JP 6**:DNA **4**)	30.8	79.3	-1.5	+28.8

Mismatched duplexes of hetero Janus PNAs JP 5 and JP 6. To further confirm the complex formation through sequence specific base pairing, duplexes from **JP 6** with DNA sequences mismatched from each face was studied. The amide face mismatched DNA **15m** (with C-T mismatch) and triazole face mismatched DNA **16m** (with A-C mismatch) were used for duplex formation with **JP 6** (Figure 4.33). The triazole face mismatch duplex **JP 6**:DNA **16m** showed a single transition corresponding T_m of 38.8 °C (Table 4.12, entry 1), which is lower by 12.7 °C than the T_m for perfect complementary duplex **JP 6**:DNA **4** (T_m = 51.5 °C). The mismatched

duplex from amide face DNA **15m:JP 6** also showed a single transition with T_m of 30.9 ºC (Table 4.12, entry 2), which is lower by 1.4 ºC than the perfect duplex.

In presence of both face mismatch DNAs (DNA **15m** and DNA **16m**), the *hetero Janus* PNA **JP 6** exhibited a single and relatively broad melting transition with T_m = 53.7 °C (Table 4.12, entry 3). Compared with their individual isolated mismatched duplexes, the T_ms of both amide and triazole duplexes in the mismatched double duplex were enhanced by +22.8 ºC and +14.9 ºC respectively. The observance of single transition in this double duplex is due to the less difference or overlap in T_ms of both duplexes. Nevertheless, enhancement of T_m similar to that seen in perfect duplexes and lowering of T_m of mismatched duplexes, substantiated formation of double duplex in *hetero Janus* PNA by sequence specific complementary base pairing.
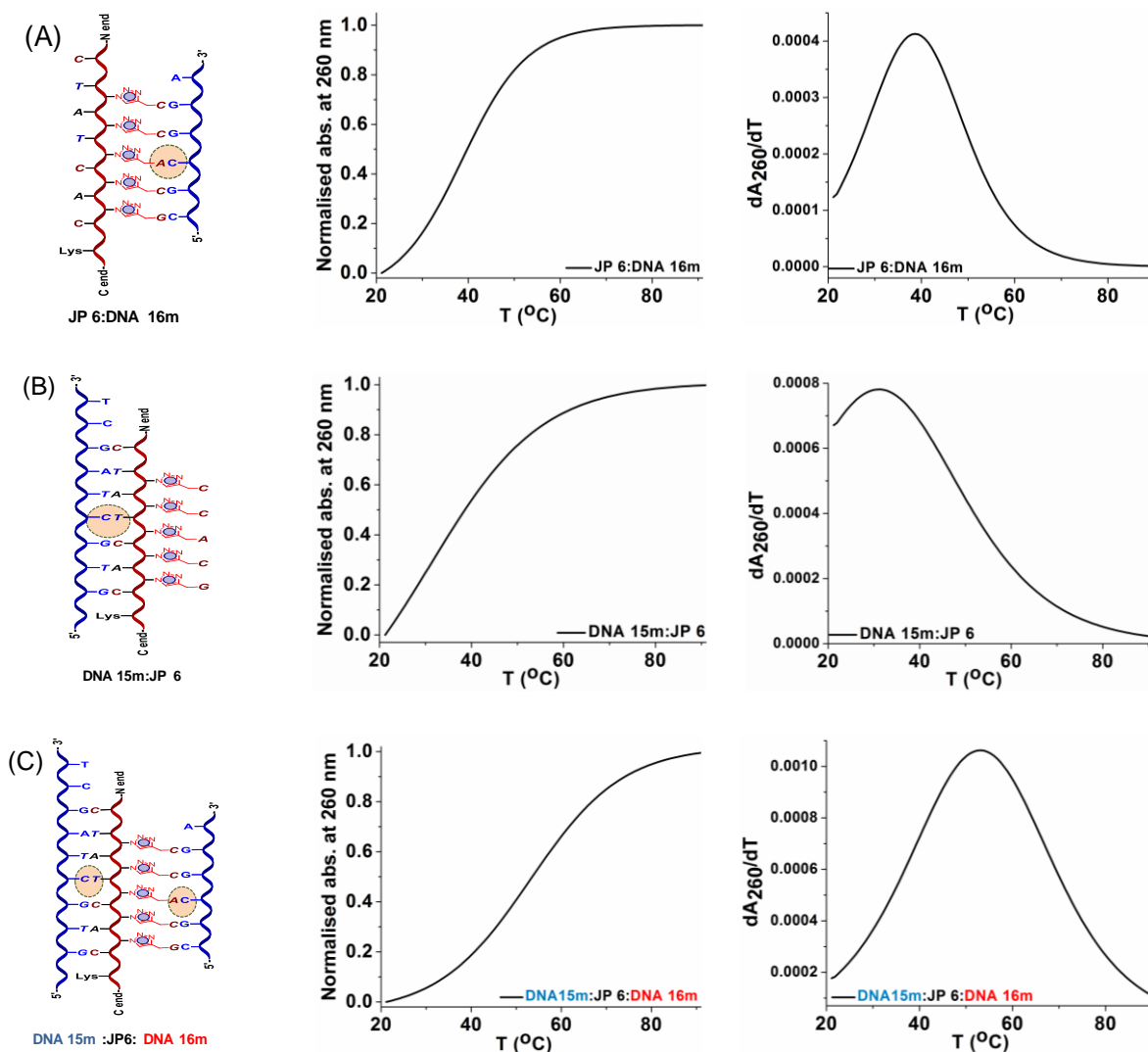
Figure 4.33 Temperature dependent UV absorbance curves for mismatch DNA duplexes of *hetero Janus* PNA **JP 6** = m*jp-tz$_5$-CCACG*, DNA **15m** = 5'-GTG**C**TAGCT-3', DNA **16m** = 5'-CG**C**GGA-3'; Buffer: 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM **(A) JP 6:DNA 16m (B)** DNA **15m**:**JP 6 (C)** DNA 15m:**JP 6**:DNA **16m**, T_m values are accurate to ±1.0 °C.

Table 4.12 UV-T_m (°C) of m*jp-tz$_5$-CCACG* (**JP 6**) with mismatch duplexes

Entry	*Hetero Janus* PNA:**mismatch**DNA complexes	T_m (°C)		ΔT_m (°C)		ΔT_m^1 (°C) (**mismatch** duplex – **perfect** duplex)	
		Amide	Triazole	Amide	Triazole	Amide	Triazole
1	**JP 6:DNA 16m**	-	38.8	-	-		-12.7
2	DNA **15m**:**JP 6**	30.9	-	-	-	-1.4	
3	DNA **15m**:**JP 6**:DNA **16m**	53.7		+22.8	+14.9		-25.6

ΔT_m^1 indicates the difference in T_m of m*jp-tz$_5$-CCACG* with mismatch DNA and cDNA.

4.4.4b CD spectra of *Hetero Janus PNAs* m*jp-tz$_4$-CACG (JP 5) and* m*jp-tz$_5$-CCACG (JP 6)* with complementary DNA 1/DNA 2

The CD spectra of amide duplex DNA **3**:m*jp-tz$_4$-CACG* (**JP 5**) shows a positive maximum at 277 nm and a negative maximum at 256 nm with cross-over points at 226 and 265 nm. The CD spectra of trizole duplex m*jp-tz$_4$-CACG*:DNA **4** shows two positive maxima at 248 nm and 281 nm and a negative maxima at 263 nm. In comparison, the double duplex DNA **3**:m*jp-tz$_4$-CACG*:DNA **4** shows positive maxima at 249 nm and 280 nm and a negative maximum at 260 nm (Figure 4.34A).

The CD spectra of *Janus* PNA with amide face duplex DNA **6**:m*jp-tz$_5$-CCACG* (**JP 6**) shows a positive maxima at 281 and a negative maxima at 256 nm with cross-over points at 238 and 263 nm,. The triazole face duplex m*jp-tz$_5$-CCACG*:DNA **4** shows a weak band with positive maxima 280 nm. The corresponding double duplex DNA **6**:m*jp-tz$_5$-CCACG*:DNA **4** shows a positive maximum at 279 nm and a negative maximum at 255 nm (Figure 4.34B).
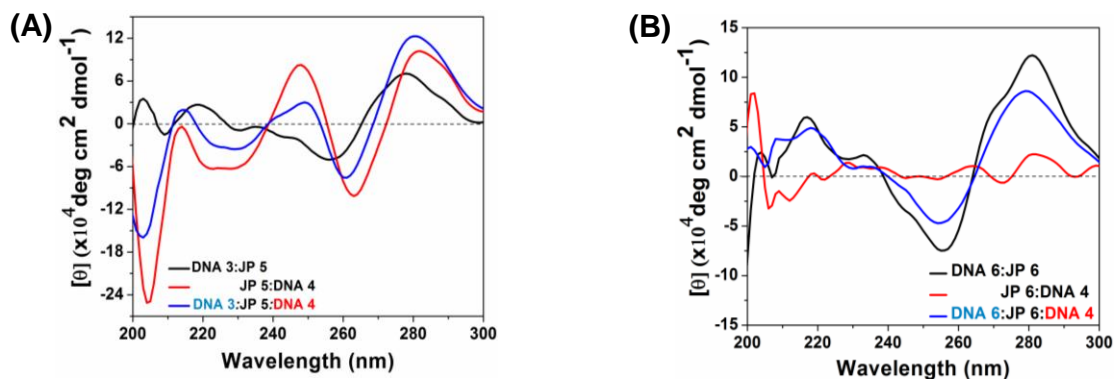
Figure 4.34 CD spectra of **(A)** *hetero Janus* PNA:DNA amide duplex DNA **3**:m*jp-tz₄-CACG* (DNA **3**:**JP 5**)**,** triazole duplex m*jp-tz₄-CACG*:DNA **4** (**JP 5**:DNA **4**) and *hetero Janus* PNA:DNA double duplex DNA 3:m*jp-tz₄-CACG*:DNA **4** (DNA **3**:**JP 5**:DNA **4**) **(B)** *hetero Janus* PNA:DNA amide duplex, DNA 6:m*jp-tz₅-CCACG* (DNA **6**:**JP 6**), m*jp-tz₅-CCACG*:DNA **4** (**JP 6**:DNA **4**). Buffer; 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM.

4.4.4c ITC studies of complexes of *mjp-tz₄-CACG* and *mjp-tz₄-CACG* with cDNA

The *hetero Janus* PNA with tetrameric triazole *mjp-tz₄-CACG*, (**JP 5,** 47.2 μM) taken in cell was titrated with amide face complementary DNA **3** (1.10 mM) to obtain the binding isotherm shown in (Figure 4.35A). The thermodynamic data extracted from the binding curve gave ΔH of -26.1 kcal/mole, K_D of 11.4 x 10^{-6} M and a binding stoichiometry N of 1:1 (N = No. of nucleobase involved in binding from DNA / no. of nucleobase involved in binding from *Janus* PNA, 5/5 = 1). This suggests the formation of 1:1 duplex of DNA **3**:**JP 5**. In the second experiment, DNA **4** (0.88 mM) complementary to triazolyl face was titrated into *mjp-tz₄-CACG* (**JP 5**, 47.2 μM) and the binding reaction gave ΔH of -33.7 kcal/mol, K_D of 8.78 x 10^{-6} M and a binding stoichiometry N of 0.82, (N = No. of nucleobase involved in binding from DNA / No. of nucleobase involved in binding from *Janus* PNA, 4/4 = 1), indicating the formation of 1:1 duplex of **JP 5**:DNA **4** as shown in (Figure 4.35B).
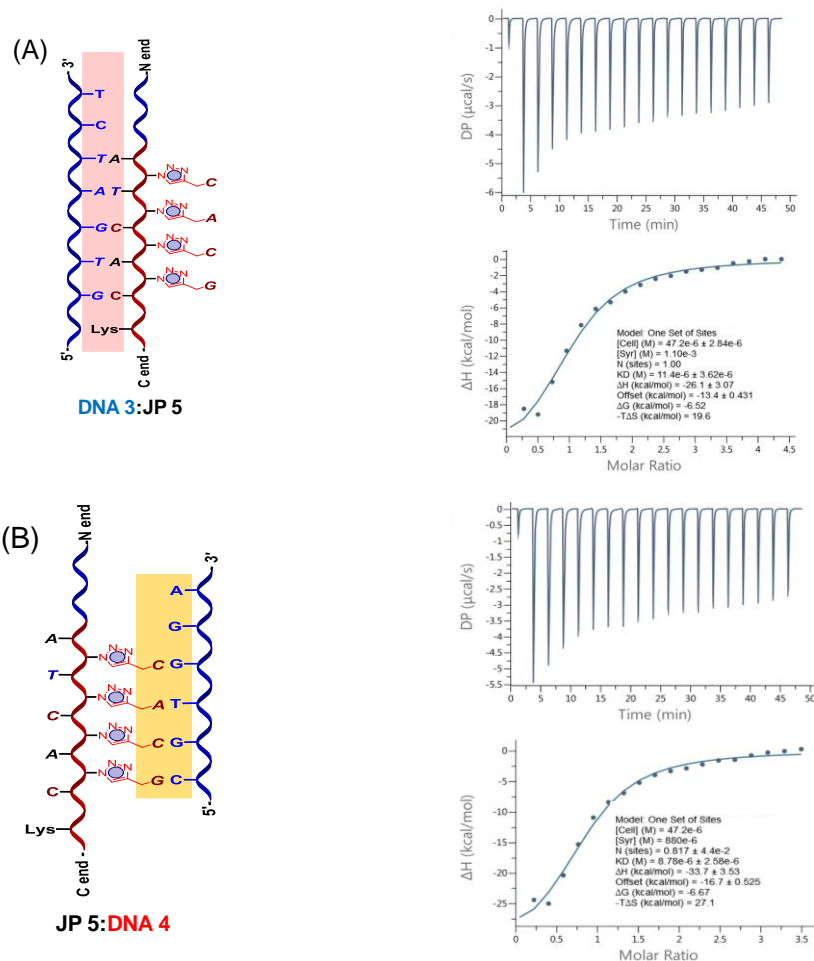
Figure 4.35 ITC data for binding of **(A)** amide face duplex DNA **3:JP 5 (B)** triazole face duplex **JP 5:**DNA **4**.

When Janus PNA *mjp-tz4-CACG* (**JP 5**, 33.5 μM) was titrated with a 1:1 (in terms of base pair from each face) mixture of DNA **3** (1.1 mM of 5-nucleobases, effective concentration per nucleobase 0.22 μM) and DNA **4** (0.88 mM of 4-nucleobase effective concentration per nucleobase 0.22 μM) that are complementary to amide and triazole faces, the binding isotherm gave ΔH of -37.9 kcal/mol, K_D of 19.5×10^{-6} M with binding stoichiometry N of 1 (N = number of nucleobases involved in binding from DNA / number of nucleobases involved in binding from *Janus* PNA, 9/9 = 1), corresponding to formation of double duplex DNA **3:JP 5:**DNA **4** (Figure 4.36).
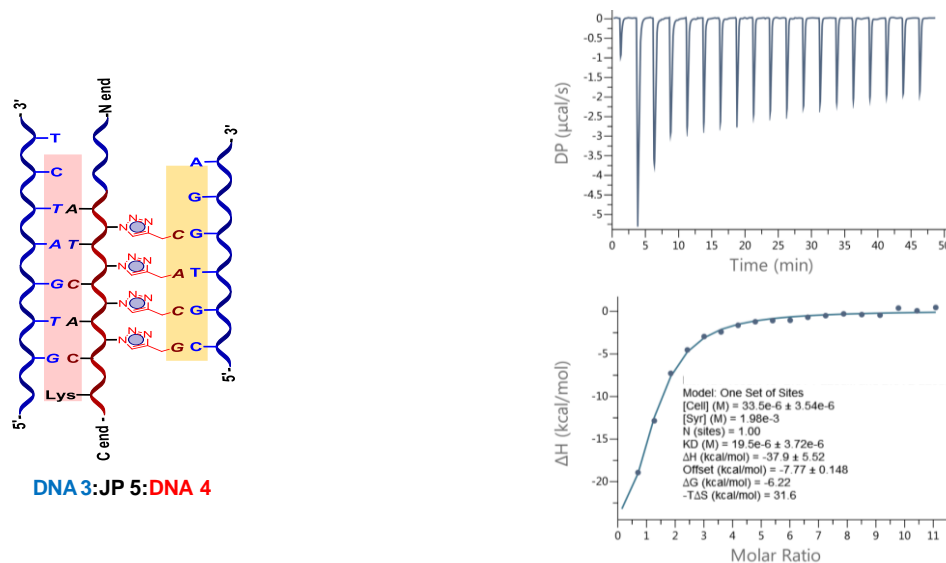
Figure 4.36 ITC data for binding of **JP 5** to mix of (DNA **3** + DNA **4**)

Binding of JP 6 to complementary DNA 4 and DNA 6 to form single duplex. The *Janus* PNA pentameric tetrazole *mjp-tz₅-CCACG* (**JP 6**, 9.82 μM) was taken in cell and titrated with amide face complementary DNA **6** (280 μM) the binding isotherm gave ΔH of -31.3 kcal/mol, K_D of 2.11×10^{-6} M and N of 1.0 (N = number of nucleobases involved in binding of DNA / number of nucleobases involved in binding from *Janus* PNA) indicating formation of DNA **6**:**JP 6** (1:1) duplex (Figure 4.37A). In a second experiment, titration of **JP 6** (9.82 μM) with triazole face DNA **4** (220 μM) gave ΔH of –39.3 kcal/mol and K_D of 7.92×10^{-6} M with N = 1.0 (in terms of No. of nucleobases involved in binding from DNA / No. of nucleobases involved in binding from Janus PNA) indicating the formation of **JP 6:**DNA **4** (1:1) duplex (Figure 4.37B).

215
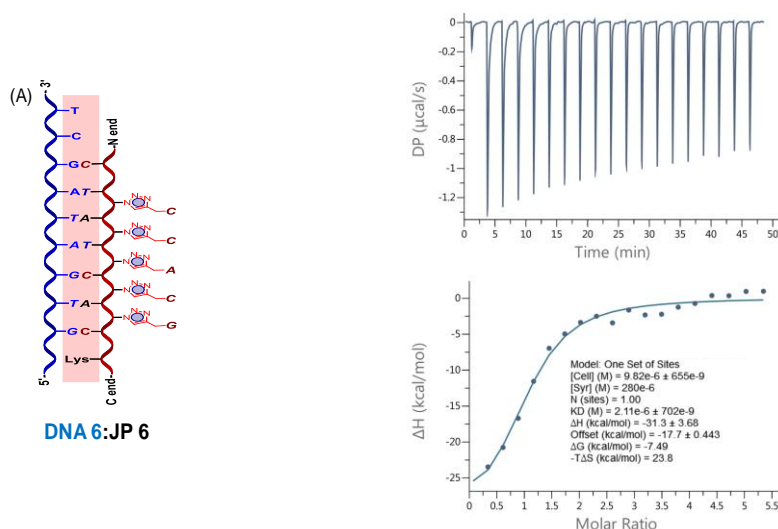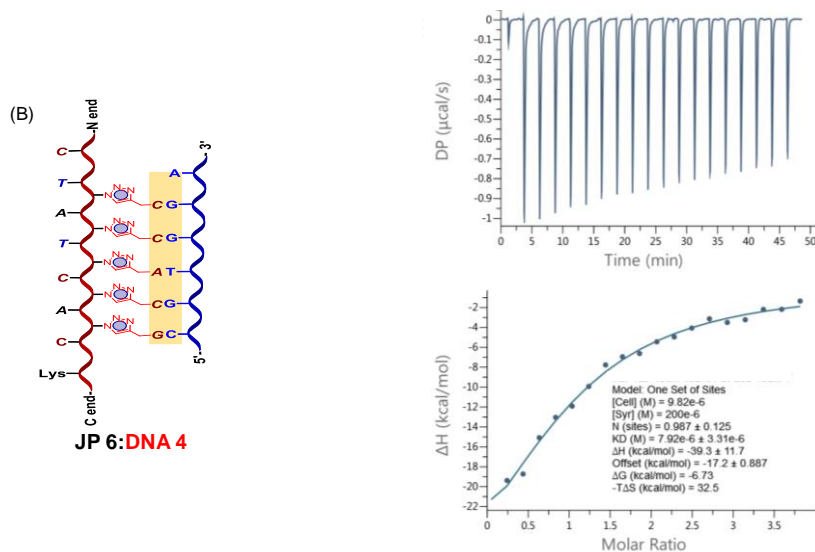
Figure 4.37 ITC data for binding of **(A) JP 6** to DNA **6 (B) JP 6** to DNA **4**

Binding of JP 6 to complementary DNA 4 and DNA 6 simultaneously to form double duplex. When *mjp-tz₅-CCACG* (**JP 6,** 15.3 μM) was titrated with a (1:1) mixture of DNA **6** (0.21 mM of 7-nucleobase, effective concentration/nucleobase 30 μM) + DNA **4** (0.15 mM of 5-nucleobase effective concentration/nucleobase 30 μM) that are complementary to amide and triazole faces respectively, the binding isotherm lead to ΔH of -20.4 kcal/mol, K_D of 2.02×10^{-6} M and N of 1.0 was obtained, indicating formation of double duplex structure (Figure 4.38).
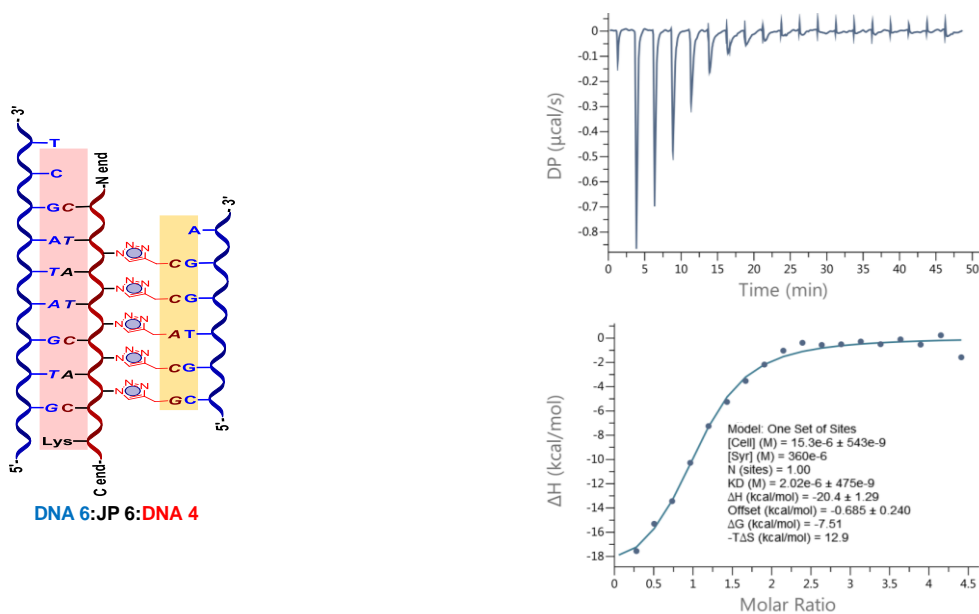
Figure 4.38 ITC data for binding of **JP 6** to mixture of (DNA **6** + DNA **4**)

Table 4.13 summarizes the thermodynamic parameters derived from ITC data in the titration of *Janus* PNAs **JP 5** and **JP 6** with amide face complementary DNA 3 and DNA 6 respectively and triazole face complementary DNA **4.**

Table 4.13 ITC summary of binding **JP 5** and **JP 6** to complementary DNA

Entry	*Hetero Janus* PNA:DNA	ΔG kcal/mol		ΔH kcal/mol		ΔS cal/mol/K		K_D M (x10^{-6})		N
		Amide	**Triazole**	**Amide**	**Triazole**	**Amide**	**Triazole**	**Amide**	**Triazole**	
1	DNA **3**:**JP 5**	-6.52 ± 0.38	-	-26.1 ± 3.1	-	-68.0 ± 1.31	-	11.4 ± 3.62	-	1.1
2	DNA **6**:**JP 6**	-7.49 ± 0.4	-	-31.3 ± 3.68	-	-82.6 ± 1.37	-	2.11 ± 0.7	-	1.0
3	**JP 5**:DNA **4**	-	-6.67 ± 0.35	-	-33.7 ± 3.53	-	-94.0 ± 1.2	-	8.78 ± 2.58	0.82
4	**JP 6**:DNA **4**	**-**	-6.73 ± 0.51	-	-39.3 ± 11.7	-	-112.8 ± 1.77	-	7.92 ± 3.31	1.0
5	DNA **3**:**JP 5**:DNA **4**	-6.22 ± 0.22		-37.9 ± 5.52		-109.7 ± 0.77		19.5 ± 3.72		1.0
6	DNA **6**:**JP 6**:DNA **4**	-7.51 ± 0.27		-20.4 ± 1.29		-44.8 ± 0.95		2.09 ± 0.47		1.0

N = No. of nucleobase involved in binding from DNA / no. of nucleobase involved in binding from *Janus* PNA

4.4.5 Binding Studies *on hetero Janus* PNAs JP 5 and JP 6 with *hairpin* DNA

4.4.5a Thermal stability of double duplexes of *hetero Janus* PNA with *hairpin* DNA

Formation of double duplexes from *Janus* PNAs as described before is a termolecular reaction and hence unfavorable in entropic terms. Reducing the association to a bimolecular reaction with defined stoichiometry will lower the unfavorable entropic factors. To achieve this, *Janus* PNA:DNA duplexes were constituted using hairpin DNA $\mathbf{5_{hp}}$ and DNA $\mathbf{7_{hp}}$ having complementary stretches on amide and triazole faces linked by a spacer sequence of 5 bases (TTTTC). The hairpin DNA $\mathbf{5_{hp}}$ and DNA $\mathbf{7_{hp}}$ were used for duplex formation with *Janus* PNA m*jp-tz₄-CACG* (**JP 5**) and m*jp-tz₅-CCACG* (**JP 6**) respectively (Figure 4.39). The tetrameric triazole Janus PNA m*jp-tz₄-CACG* (**JP 5**) gave two transitions (Figure 4.39A), one for the amide face duplex at T_m of 34.6 °C and another at T_m of 65.4 °C for the triazole face duplex. It may be mentioned that the corresponding open chain double duplexes gave only one transition (Figure 4.31C). The T_ms of two transitions seen with hairpin DNA $\mathbf{5_{hp}}$:**JP 5**:DNA $\mathbf{5_{hp}}$ duplexes

are higher by 8.8 °C and 30.8 °C (for triazole and amide face respectively) than that observed with *non-hairpin* (open) DNA.

The double duplexes from pentameric triazole *hetero Janus* PNA m*jp-tz₅-CCACG* (**JP 6**) with DNA **7ₕₚ** were also stabilized with T_ms of 42.5 °C for amide face and 78.8 °C for triazole face (Figure 4.39B), compared with that of corresponding open chain double duplexes ΔT_m +11.7 °C for amide face and − 0.5 °C for triazole face (Figure 4.32C). Thus all features of stability of double duplexes seen for *non-hairpin Janus* PNA:DNA duplexes namely (i) triazole face duplex having higher stability than amide face duplex and (ii) individual separate transitions seen for amide face and triazole face duplexes in double duplexes and (iii) synergistic enhancement in the stability of both face duplexes, were all retained in DNA hairpin duplexes of *Janus* PNA as well. The observed relative stability of duplexes of various *hetero Janus* PNA:DNA duplexes are summarized in Table 4.14.
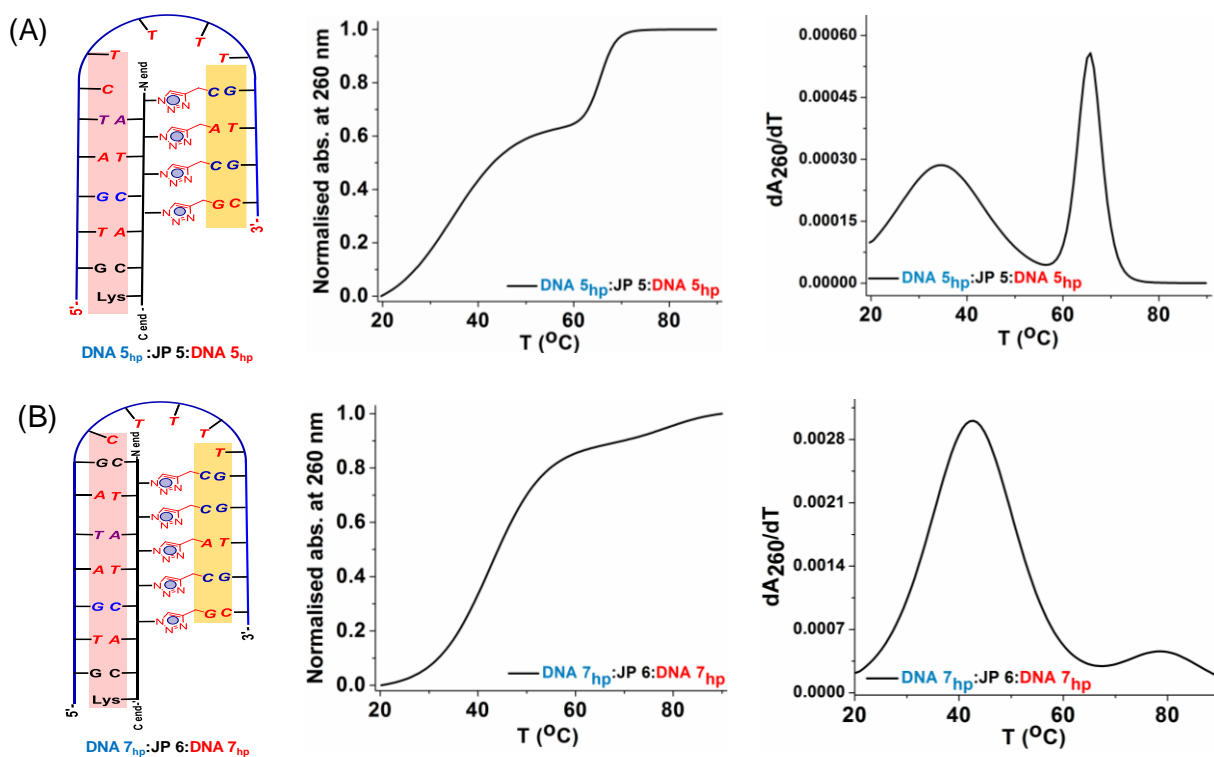
Figure 4.39 Temperature dependent UV absorbance curves for duplexes of *hetero Janus* PNAs m*jp-tz₄-CACG* (**JP 5**) and m*jp-tz₅-CCACG* (**JP 6**) with complementary hairpin DNA duplexes. DNA **5ₕₚ** = 5'-GTGATCTTTTGTGC-3'; DNA **7ₕₚ** = 5'-GTGATAGCTTTTGGTGC-3'; Buffer: 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM) **(A)** DNA **5ₕₚ:JP 5:**DNA **5ₕₚ**. **(B)** DNA **7ₕₚ:JP 6:**DNA **7ₕₚ**, T_m values are accurate to ± 1.0 °C.

Table 4.14 UV-T_m (^oC) of *hetero Janus* PNA m*jp-tz₄-CACG* (**JP 5**) and m*jp-tz₅-CCACG* (**JP 6**):DNA with open vs hairpin DNA duplexes

Entry	*Hetero Janus* **PNA:DNA complexes (open vs hairpin)**	T_m (^oC)		ΔT_m (^oC) hairpin – open double duplex	
		Amide	Triazole	Amide	Triazole
1	DNA 5$_{hp}$:**JP 5**:DNA 5$_{hp}$	34.6	65.4	-22.0	+8.8
2	DNA 3:**JP 5**:DNA 4	56.6		-	-
3	DNA 7$_{hp}$:**JP 6**:DNA 7$_{hp}$	42.5	78.8	+11.7	-0.5
4	DNA 6:**JP 6**:DNA 4	30.8	79.3	-	-

ΔT_m (hairpin-open) indicates the difference in T_m imposed in hairpin duplexes compared open chain duplexes.

4.4.5b CD spectra of complexes of *Hetero Janus PNAs* m*jp-tz₄-CACG* and m*jp-tz₅-CCACG* with complementary DNA (DNA 3/DNA 4/DNA 6) and hairpin DNA (DNA 5$_{hp}$/DNA 7$_{hp}$)

The CD spectra of double duplexes from tetrameric tetrazole *Janus* PNA m*jp-tz₄-CACG* (**JP 5**) *and* pentameric tetrazole m*jp-tz₅-CCACG* (**JP 6**) with respective complementary DNA hairpin loop DNA **5$_{hp}$** and DNA **7$_{hp}$** show a positive maximum at 279 nm and a negative maximum around 254 nm with cross-over points at 227 and 263 nm**.** Comparative CD spectra of *hetero Janus* PNA (**JP 5** / **JP 6**):DNA complexes with open DNAs DNA **3,** DNA **4,** DNA **6** vs hairpin DNAs DNA **5$_{hp}$,** DNA **7$_{hp}$** are shown in (Figure 4.40) and the characteristic CD band pattern of duplexes summarized (Table 4.15). All show a somewhat similar pattern with positive maxima around 280 nm and negative maxima at 250 nm. Only the double duplex of **DNA 3:JP 5:DNA 4** showed slight variable shifts of both positive and negative bands to longer wavelengths by 5 nm. Since the CD bands originate primarily from the DNA component, not much changes are seen for various Janus PNA complexes.
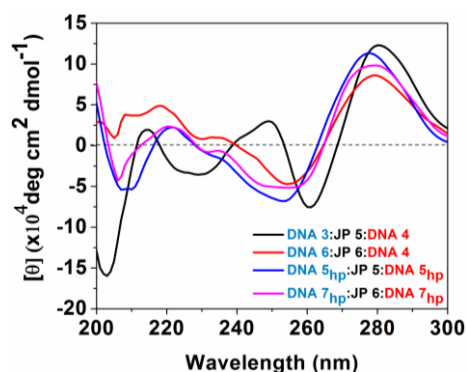
Figure 4.40 CD spectra of open, DNA **3**:JP **5**:DNA **4** (DNA 3:m*jp-tz₄*-*CACG*:DNA 4), DNA **6**:**JP 6**:DNA **4** (DNA 6:m*jp-tz₅*-*CCACG*:DNA 4) vs hairpin, DNA **5**ₕₚ:**JP 5**:DNA **5**ₕₚ (DNA 5ₕₚ:m*jp-tz₄*-*CACG*:DNA 5ₕₚ) and DNA **7**ₕₚ:**JP 6**:DNA **7**ₕₚ (DNA 7ₕₚ:m*jp-tz₅*-*CCACG*:DNA 7ₕₚ) double duplex complex Buffer: 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM.

Table 4.15 Summary of CD band pattern of *hetero Janus* PNAs complex

Entry	m*jp-tz₄*-*CACG* (**JP 5**) complex	+ ve band (nm)	- ve band (nm)	m*jp-tz₅*-*CCACG* (**JP 6**) complex complex	+ ve band (nm)	- ve band (nm)
1	DNA **3**:**JP 5** (Amide duplex)	277	256	DNA **6**:**JP 6** (Amide duplex)	281	256
2	**JP 5**:DNA **4** (Triazole duplex)	279	262	**JP 6**:DNA **4** (Triazole duplex)	280	255 (weak band)
3	DNA **3**:**JP 5**:DNA **4** (double duplex)	280	260	DNA **6**:**JP 6**:DNA **4** (double duplex)	279	255
4	DNA **5**ₕₚ:**JP 5**:DNA **5**ₕₚ (double duplex)	277	251	DNA **7**ₕₚ:**JP 6**:DNA **7**ₕₚ (double duplex)	279	260

4.4.5c ITC studies of binding of *Hetero Janus* PNAs *JP 5* and *JP 6* to hairpin DNAs

This section describes the ITC experiments of binding of *Hetero Janus* PNAs m*jp-tz₄*-*CACG* (**JP 5**) and m*jp-tz₅*-*CCACG* (**JP 6**) to complementary DNA **5**ₕₚ and DNA **7**ₕₚ respectively. **JP 5** (60.2 μM) was titrated with hairpin DNA **5**ₕₚ (1.98 mM) and the ITC data showed single sigmoidal curve and gave binding curve with ΔH of -24.8 kcal/mol and K_D value of 2.11 x 10^{-6} M (Figure 4.41A). Similarly **JP 6** (28.3 μM) titrated with hairpin DNA **7**ₕₚ (1.2 mM) also gave a single transition with ΔH of -21.7 kcal/mol and K_D value 0.61x10^{-6} M (Figure 4.41B), binding **JP 5** and **JP 6** to **DNA**ₕₚ summarises in Table 4.16.

(A)

DNA **5**ₕₚ:**JP 5**:DNA **5**ₕₚ

Model: One Set of Sites
[Cell] (M) = 60.2e-6 ± 1.56e-6
[Syr] (M) = 1.98e-3
N (sites) = 1.00
KD (M) = 2.11e-6 ± 589e-9
ΔH (kcal/mol) = -24.8 ± 1.00
Offset (kcal/mol) = -7.58 ± 0.224
ΔG (kcal/mol) = -7.49
-TΔS (kcal/mol) = 17.3
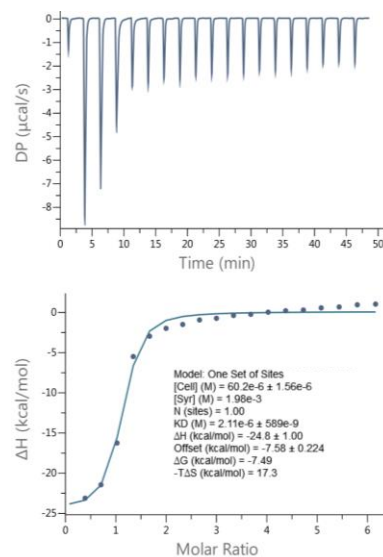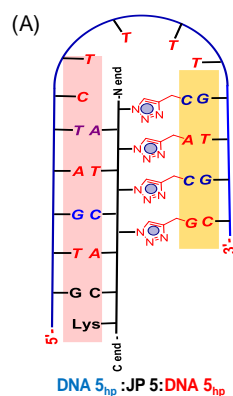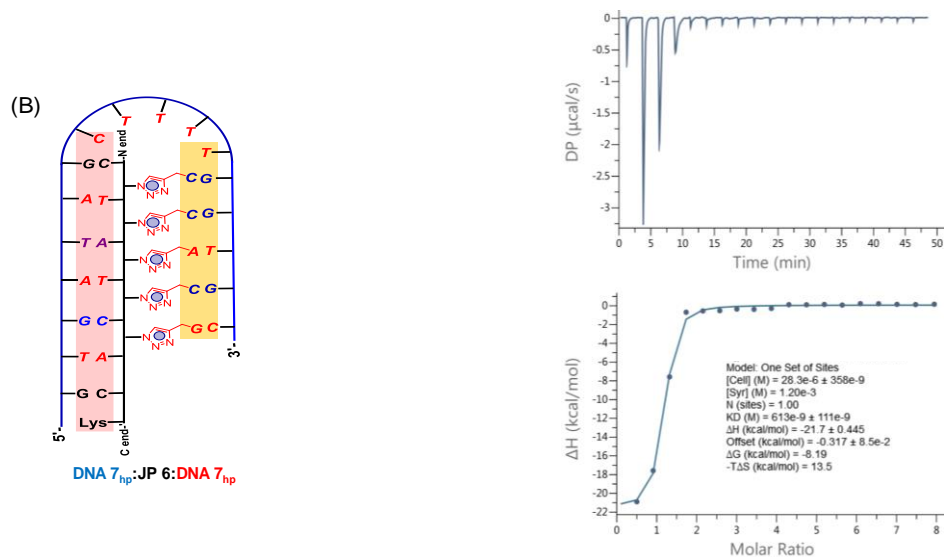
Figure 4.41 ITC data for binding of *Hetero Janus* PNA with hairpin loop DNA **(A) JP 5** to DNA **5_{hp}** **(B) JP 6** to DNA **7_{hp}**

Table 4.16 ITC summary of binding **JP 5** and **JP 6** to **hairpin DNA**

Entry	*Hetero Janus* PNA:DNA$_{hp}$	ΔG kcal/mol	ΔH kcal/mol	ΔS cal/mol/K	K_D M ($\times 10^{-6}$)	N_b
1	DNA 5$_{hp}$:**JP 5**:DNA 5$_{hp}$	-7.49 ± 0.33	-24.8 ± 1.0	-60.0 ± 1.14	2.11 ± 0.59	1.0
2	DNA 7$_{hp}$:**JP 6**:DNA 7$_{hp}$	-8.19 ± 0.21	-21.7 ± 0.44	-46.8 ± 0.73	0.61 ± 0.11	1.0

N_b = No. of nucleobase involved in binding from DNA / no. of nucleobase involved in binding from *Janus* PNA

4.4.6 Binding studies of *Self Complementary Mix Janus* PNA (*SCM-JP*) with complementary DNA

The sequences of *Janus* PNAs on amide and triazole faces described so far are non-complementary to each other. The *self-complementary Janus* PNAs have potential to grow into 2D assembled structures by continuous intermolecular H-bonding of Janus PNA chains. The duplex forming ability of such self-complementary sequences were studied by UV and CD techniques.

4.4.6a Thermal stability of DNA duplexes from *self-complementary mix Janus* PNA (*SCM-JP*)

In *Janus* PNA **JP 7** (*SCM-JP*), the sequences of the amide and triazole faces are complementary to each other. The mixed 5-mer base sequence on triazole face is complement to

a stretch within the 7-mer base sequence on amide face. This will enable Janus PNA both self-complementation and formation of double duplexes with complementary DNA as in previous examples.

The UV-T_m profiles of DNA duplexes of the *self-complementary mix Janus* PNA **JP 7** (*SCM-JP*) with cDNA **8** (amide face) and cDNA **9** (triazole face) are shown in (Figure 4.42). Both the amide duplex DNA **8:JP 7** and the *triazole duplex* **JP 7:DNA 9** exhibited single transitions corresponding to T_ms of 34.9 ºC and 51.6 ºC respectively. Again the triazole duplex has higher thermal stability than the amide side duplex, inspite of having a shorter sequence. The double duplex from mixed sequence *self-complementary Janus* PNA DNA **8:SCM-JP 7:DNA 9** displayed two transitions with $T_{m1} = 46.0$ °C for amide face duplex and $T_{m2} = 82.1$ °C for triazole face duplex. The thermal stability of both amide and triazole side duplexes was enhanced in double duplex, compared to single duplex alone by +11.1 °C and +30.5 °C respectively (Table 4.17, entry 3). As in previous cases, the triazolyl-face showed higher stabilization than amide face by +16.7 ºC in mono duplex form and +36.1 ºC in double duplex state.
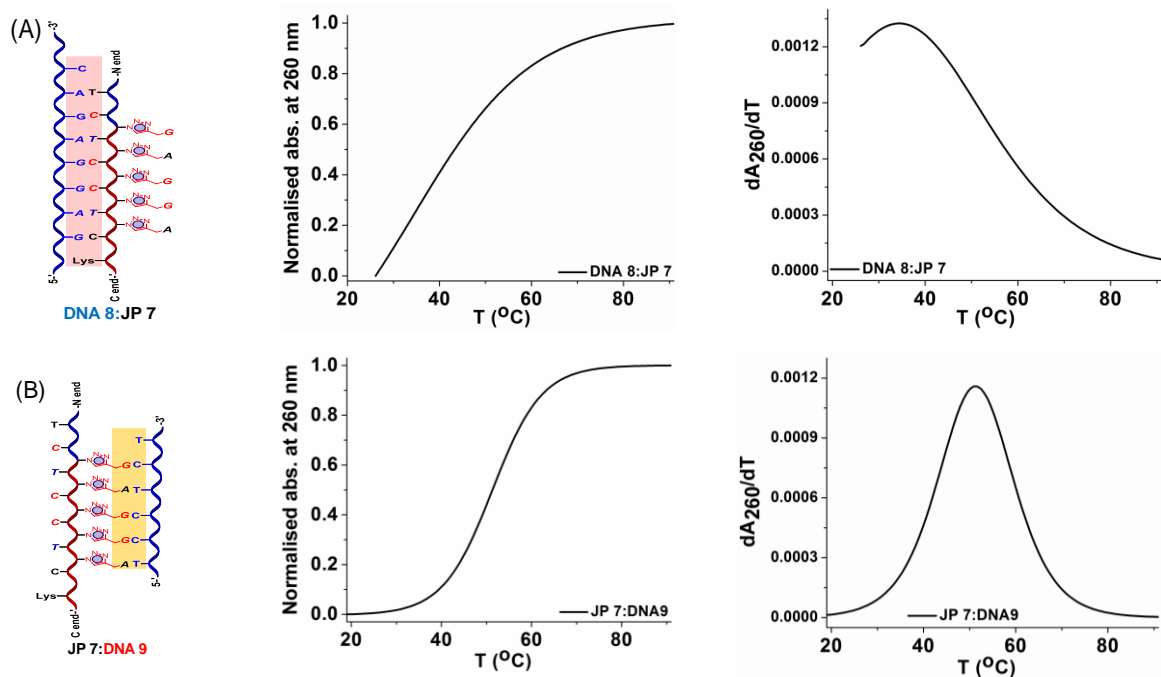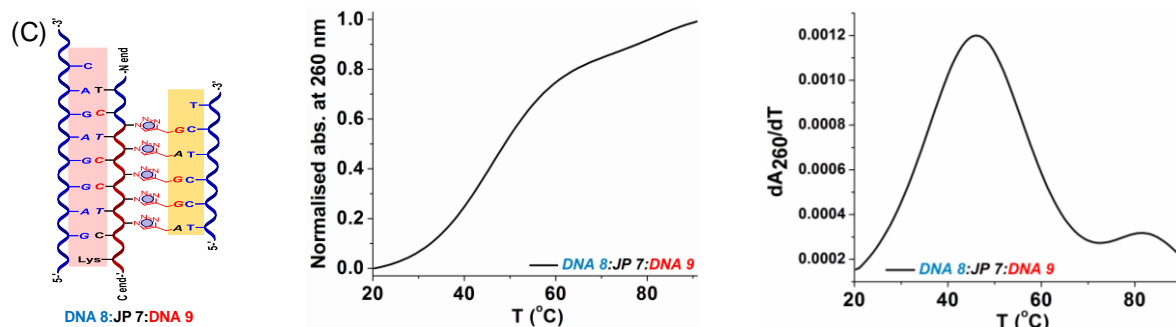
Figure 4.42 Temperature dependent UV absorbance curves for *Self Complementary hetero Janus*-PNA **JP 7** (*SCM-JP*):DNA duplexes; DNA **8** = 5'-GAGGAGAC-3', DNA **9** = 5'-TCCTCT-3' Buffer: 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM. **(A)** DNA 8:*SCM-JP* (DNA **8:JP 7**) **(B)** *SCM-JP*:DNA 9 (**JP 7**:DNA **9**) **(C)** DNA 8:*SCM-JP*:DNA 9 (DNA **8:JP 7**:DNA **9**), T_m values are accurate to ±1.0 °C.

Table 4.17 UV-T_m (°C) of *self complementary mix Janus-PNA* (**JP 7** or *SCM-JP*):DNA duplexes

Entry	*Self complementary mix Janus* **PNA:DNA complexes**	T_m (°C)		ΔT_m (°C) (double duplex - T_m duplex)	
		Amide	Triazole	Amide	Triazole
1	**JP 7**:DNA **9** (*SCM-JP*:DNA **9**)	-	51.6	-	-
2	DNA **8**: **JP 7** (DNA **8**:*SCM-JP*)	34.9	-	-	-
3	DNA **8:JP 7**:DNA **9** (DNA **8**:*SCM-JP*:DNA **9**)	46.0	82.1	+11.1	+30.5

Concentration-dependent UV-T_m of self-complementary mix Janus PNA (JP 7). The *self-complementary mix Janus* PNA *SCM-JP* (**JP 7**) can also form PNA:PNA duplex and in principle antiparallel orientation by self-hybridization of amide to triazole face binding is as shown in (Figure 4.43). Since self-association is expected to increase with concentration, the thermal stability of self-duplex from *SCM-JP* (**JP 7**) was examined as a function of concentration from 1 µM to 5 µM. The *Janus SCM-JP* (**JP 7**) on its own formed self duplex as seen by a nice sigmoidal transition. With increase in concentration of *SCM-JP* (**JP 7**)**,** the T_m showed increase from 47 °C to 53.9 °C, with a jump of 4 °C in the T_m from 3 µM to 4 µM step (Figure 4.43). In this type of self-duplex formation, there are no separate amide / triazole face and hence only one transition is observed.
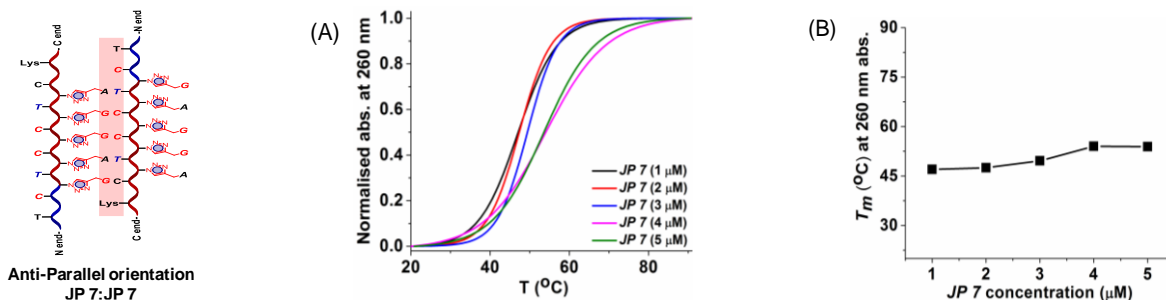
Figure 4.43 Concentration-dependent UV-T_m curves for *Self-Complementary Mix Janus-PNA or SCM-JP* (**JP 7**) **(A)** Normalised absorbance at 260 nm **(B)** T_m value (°C) vs concentration in (μM).

4.5 Discussion

The temperature dependent UV-spectroscopic studies determine the thermal stability of *Janus* PNA oligomers with complementary DNA. The UV-*T* and CD data of DNA duplexes derived from the new backbone *p*7-*tz*-*C5/G5* clearly suggested that nucleobases linked at C_α through a triazole linker on its own can from duplexes by sequence specific base pairing, similar to that derived from well-established standard *aeg*-PNA backbone. Further, the G bases linked to triazole backbone give higher stability than analogous C-bases. The UV-T_m results compiled for all Janus PNA oligomers in (Table 4.18) showed that the Janus PNAs having nucleobases on amide and triazole faces can form independent duplexes from each face with corresponding cDNA and show expected single transition (entry 1, 2, 4, 5, 7, 8, 10, 11, 13 and 14). The T_m of triazole side duplexes were always higher than those from amide duplexes, with ΔT_m (T_m triazole − T_m amide) ranging from +16.7 °C to +20.5 °C. Significantly, Janus PNAs form duplexes with appropriate complementary DNA on both faces leading to double duplexes or double duplex of triplex, depending on the nature of sequences. This is reflected in observance of two separate transitions, corresponding to duplex and triplex (Figure 4.13C) or two duplexes (Figure 4.32C). In *Janus* PNA double duplexes, two transitions were noticed with T_m of triazole duplexes always higher than T_m of amide duplexes (entry 3, 6, 12 and 15), with ΔT_m (T_m triazole − T_m amide) ranging from +27.7 °C to +48.5 °C. This trend was consistent with all Janus PNAs studied: *homo Janus* PNA, *chimeric Janus* PNA and *hetero Janus* PNA. The two single stranded complementary DNAs for the two faces were linked by a spacer chain to form hairpin DNA:*Janus* duplex (Figure 4.39) and the results were similar to open chain duplexes, with T_m of duplex from triazole face higher than that of duplex from amide face.

The successful and simultaneous formation of duplexes from both sides by complementary base pairing was supported by lowered T_m in duplexes with mismatched bases in cDNA strand on either amide or triazole and in both sides, which is very similar to reported decrease in T_m of *aeg*PNA with mismatch DNA duplex.[19]

Table 4.18 Summary of UV-T_m data

Entry	*Janus* PNA:DNA complexes	T_m (ºC)		T_m (double duplex) - T_m duplex		T_m (Triazole) -T_m (Amide) (ºC)
		Amide	Triazole	Amide	Triazole	
1	T_7jp-tz-C_5:dG$_6$ (**JP 1**:DNA **2**) Triazole duplex	-	50.7			+20.5
2	C_5-tz-jpT_7:dA$_8$:T_7jp-tz-C_5 (**JP1**:DNA **1**:**JP 1**) Amide triplex	30.2	-	-		
3	dG$_6$:C_5-tz-jpT_7:dA$_8$:T_7jp-tz-C_5:dG$_6$ (DNA **2**:JP **1**:DNA **1**:JP **1**:DNA **2**) double duplex of triplex	45.8	73.5	+15.6	+22.8	+27.7
4	mjp-tz-C_5:dG$_6$ (**JP 2**:DNA **2**) Triazole duplex	-	51.2			+17.6
5	cDNA:mjp-tz-C_5 (DNA **3**:**JP 2**) Amide duplex	33.6	-	-		
6	cDNA:mjp-tz-C_5:dG$_6$ (DNA **3**:JP **2**:DNA **2**) double duplex	38.8	77.1	+5.2	+25.9	+38.3
7	mjp-tz$_4$-*CACG*:DNA 4 (**JP 5**:DNA **4**) Triazole duplex	-	48.7			+18.5
8	DNA **3**:mjp-tz$_4$-*CACG* (DNA **3**:JP **5**) Amide duplex	30.2	-	-		
9	DNA **3**:mjp-tz$_4$-*CACG*:DNA **4** (DNA **3**:JP **5**:DNA **4**) double duplex	56.6		+26.4	+7.9	-
10	mjp-tz$_5$-*CCACG*:DNA **4** (**JP 6**:DNA **4**) Triazole duplex	-	51.5			+19.2
11	DNA **6**:mjp-tz$_5$-*CCACG* (DNA **6**:JP **6**) Amide duplex	32.3	-	-		
12	DNA **6**:mjp-tz$_5$-*CCACG*:DNA **4** (DNA **6**:JP **6**:DNA **4**) double duplex	30.8	79.3	-1.5	+28.8	+48.5
13	*SCM*-*JP*:DNA **9** (**JP 7**:DNA **9**) Triazole duplex	-	51.6			+16.7
14	DNA **8**:*SCM*-*JP* (DNA **8**:JP **7**) Amide duplex	34.9	-			

15	DNA **8**:*SCM*-*JP*:DNA **9** (DNA **8**:**JP 7**:DNA **9**) double duplex	46.0	82.1	+11.1	+30.5	+36.1

The most interesting aspect of Janus PNAs is that the T_m of each duplex (from amide face or triazole face) in double duplex from Janus PNAs were enhanced significantly over that of isolated duplexes from individual faces, with ΔT_m ranging from +5.7 ^{o}C to +30.5 ^{o}C. *This is an unexpected but highly interesting and significant outcome of the present work. Thus, the most notable feature of double duplex formation by Janus PNAs is the synergistic enhancement in stability observed for each face duplex from the formation of duplex from the other face*

The *Janus* PNAs themselves alone in single stranded form have very weak CD profiles, inspite of carrying a chiral center at C_α of glycine segment. This is perhaps due to the fact that the base chromophores are far away from the chiral center to get any chiral induction. But upon complementation with DNA, the duplexes show defined CD profiles. The amide face and triazole face duplexes show individual characteristic CD patterns, the triazole face duplexes showing higher intensity CD bands around 270 nm, compared to amide face PNA:DNA duplex CD bands b/w 220 − 230 nm. (Figure 4.15, 4.28 and 4.34). The study of order of duplex formation of double duplexes i.e., amide face first followed by triazole face or vice versa, by changing the sequence of complementation (Figure 4.18), indicated that it does not matter and the resulting double duplex always attains the same final conformation.

The *triazole PNA* wherein the nucleobases are anchored to C_α - side chain through triazole linker (*p*7-*tz*-*C₅*/*G₅*) (devoid of tertiary amide bases) formed nice duplexes showing single transition (Figure 4.8 and 4.9). This is a new structural class of PNA analogues not known till now. It is known from literature[20] on various PNA analogues that the restricted rotation imparted by tertiary amide linkage in aminoethylglycyl PNA backbone is structurally and geometrically important for PNA to form strong hybrids with DNA. The replacement of carbonyl of tertiary amide by methylene destabilized the derived duplexes.[21] In this context, it is noteworthy that the new class of triazole PNAs reported here, that lack the tertiary amide group and having nucleobases linked by an alkyl spacer containing triazole moiety still show good duplex formation, with stability as good as that of standard *aeg*-PNA. The ability and stability of duplex from these triazole PNAs suggest that triazole stacking may have an important role in

addition to canonical base pairing and stacking of complementary base pairs. The additional triazole stacking is also responsible for enhanced duplex formation in *Janus* PNAs.

In order to understand the driving force for double duplex formation and the synergistic effects on observed mutual stabilities, thermodynamic parameters for various duplex formation were evaluated using ITC (Summary Table 4.19). The experiments were done by taking stoichiometric amounts of cDNA in syringe and titrating against the PNA samples taken in the cell. The titration curves showed good sigmoidal profiles (Figures 4.22 - 4.25, 4.35 - 4.38 and 4.41) and from these data the dissociation constant (K_D), stoichiometry of binding in terms of ratio of number of base pairs (N), the enthalpy (ΔH), entropy (ΔS) and free energy (ΔG) for various binding reactions were evaluated.

The dissociation constant K_D values suggest that in *homo Janus* PNA complexes, the triplexes and double duplexes have a higher binding (lower K_D) compared to simple duplexes. (entry 1-4 and entry 6-9). Formation of hairpin Janus duplex of duplex increases the binding (lower K_D) compared to corresponding open chain duplexes (entry 10-13).

The ITC data for *homo Janus* PNA **JP 1** (T_7jp-tz-C_5) titrated with a mixture of dA_8+dG_6 (1:2) ratio show formation of triplex from amide T_7 face with dA_8 (T_7:A_8:T_7) and duplex from triazole C_5 face (C_5:G_6) leading to dG_6:C_5-tz-jpT_7:dA_8:T_7jp-tz-C_5:dG_6, that is duplex (blue text) of a triplex (red text).

The stoichiometry of binding N (N = No. of nucleobases involved in binding from DNA / No. of nucleobase involved in binding from *Janus* PNA), computed from ITC data in terms of number of base pairs (*not molar concentrations*) nicely corresponded to the composition of envisaged duplexes. In all cases one set of sites binding model was used for fitting data, except for entry 3, which had two sets of sites binding model. For all duplexes, N was very close to 1.0 (entry 2, 4-11).

The enthalpy ΔH of duplexes from each face in individual *Janus* PNA complexes are in the range -20.0 to -40.0 kcal/mol, except for *homo Janus* PNA duplex **JP 1**:DNA **2**, which was slightly lower. The enthalpies of the amide duplexes are relatively higher compared to that of triazole duplexes (entry 1-3). The overall range of enthalpy seen are similar to that of PNA:DNA duplexes.[22-28] The *hairpin double duplexes* in one set of sites binding mode and the corresponding open chain *hetero Janus* PNA open duplexes showed single ITC transition and

hence only one enthalpy (entry 8-11) for formation of double duplexes. However, these showed two transitions in UV-T_m experiment. In comparison with the control duplexes (entry 4-7), the triazole duplexes had higher enthalpy than the amide face duplex as also reflected in higher UV-T_m.

The entropic changes (ΔS) measured for different binding reactions were negative, indicating that these interactions were not entropically favourable. Notable exception was the duplex of triplex formed by *homo Janus* PNA where the binding of the triazole face was accompanied by entropy gain (positive entropy). In the case of *homo Janus* PNA, triplex formation at the amide face resulted in a ΔS value of -69 cal/mol/K when the triazole face was unpaired. However, in the triplex of duplex state, the ΔS of the same amide face binding showed -6.1 cal/mol/K (entry 1 and 2), which corresponds to a 10-fold increase. Although the interaction is driven by enthalpy, this data suggests that duplex formation at the triazole face is favorably linked to triplex formation at the amide face. Similarly, the entropy change in the triazole face was +1.2 cal/mol/K when the amide face was unpaired and +15.9 cal/mol/K when the amide face was engaged with its complementary strand (entry 3). Thus, triplex formation on the amide face favored the duplex formation at the triazole face through an entropy gain. Thus, binding of one face to its complementary strand preorganized the other face to bind to its complementary strand. Such a preorganization could be the underlying mechanism behind the synergistic effect observed in the thermal melting studies, where the amide face and triazole face showed an enhancement of T_m by 15.6 °C and 22.8 °C, respectively.

In case of duplexes formed by *hetero Janus* PNAs, the entropy increased with the length of the PNA. The duplex formation at the amide face of **JP 5** and **JP 6** was associated with entropy changes of -68.0 cal/mol and -82.6 cal/mol, respectively (entry 4 and 6). A similar trend was seen in the triazole side where the entropy change was -94.0 cal/mol and -112.8 cal/mol, respectively (entry 5 and 7). Increasing negative values of ΔS suggests more unfavorable entropic contribution. Thus, these duplexes were stabilized by stronger enthalpic contribution. However, comparison of ΔS of the open form of double duplexes formed by **JP 5** and **JP 6** reveals that as the length of the *Janus* PNA increased, the ΔS values became less negative (-109.7 cal/mol/K and -44.8 cal/mol/K, respectively; entry 8 and 10) suggesting that entropy change was less unfavorable when duplex formation occurs at both faces simultaneously.

Similar trend was observed when these hairpin complementary strands were bound to the hetero Janus PNAs; the ΔS values were -60.0 cal/mol/K and -46.8 cal/mol/K for **JP 5** and **JP 6**, respectively (entry 9 and 11). This effect contributes significantly to the overall ΔG changes in these *chimeric Janus* PNAs.

The free energy difference (ΔG) for all duplexes were in the range of -6 to -9 kcal/mol and no major differences were observed among individual duplexes or those in double duplexes. Thus, all the binding reactions were equally favorable. At the temperature studied (15 °C), the enthalpic contribution to free energy was slightly higher than entropic contribution in almost all complexes, suggesting that the binding reaction is perhaps driven by enthalpic considerations.

It is well known that UV-T_m and ITC data (k_d) do not match since one is done under variable temperature conditions and the other is under isothermal conditions. The kinetics of association and dissociation are different under two conditions and hence final data are not comparable.[29]

Table 4.19 Summary of ITC data

Entry	*Janus* PNA:DNA	ΔG kcal/mol		ΔH kcal/mol		ΔS cal/mol/K		K_D M (x 10^{-6})		N	UV-T_m (°C)	
		Amide	Triazole	Amide	Triazole	Amide	Triazole	Amide	Triazole		Amide	Triazole
1	**JP 1**:DNA 1:**JP 1** (triplex)	-8.52	-	-28.4	-	-69.0	-	0.34	-	0.5	30.2	-
2	**JP 1**:DNA 2 (duplex)	-	-7.51	-	-7.15	-	+1.2	-	2.0	1.08	-	50.7
3*	DNA 2:JP 1:DNA 1 JP 1:DNA 2 (Double duplex)	-8.85	-	-10.6		-6.1	-	0.2	-	0.76	45.8	73.5
		-	-7.39	-	-2.78	-	+15.9	-	2.51	1.28		
4	DNA 3:**JP 5**	-6.52	-	-26.1	-	-68.0	-	11.4	-	1.1	30.2	-
5	**JP 5**:DNA 4	-	-6.67	-	-33.7	-	-94.0	-	8.78	0.82	-	48.7
6	DNA 6:**JP 6**	-7.49	-	-31.3	-	-82.6	-	2.11	-	1.0	32.3	-
7	**JP 6**:DNA 4	-	-6.73	-	-39.3	-	-112.8	-	9.72	1.0	-	51.5
8	DNA 3:**JP 5**:DNA 4	-6.22		-37.9		-109.7		19.5		1.0	56.6	
9	DNA 5hp:**JP 5**:DNA 5hp	-7.49		-24.8		-60.0		2.11		1.0	34.6	65.4
10	DNA 6:**JP 6**:DNA 4	-7.51		-20.4		-44.8		2.09		1.0	30.8	79.3
11	DNA 7hp:**JP 6**:DNA 7hp	-8.19		-21.7		-46.8		0.61		1.0	42.5	78.8

N = No. of nucleobase involved in binding from DNA / No. of nucleobase involved in binding from *Janus* PNA.

4.6 Conclusion

Following important inferences are drawn from the biophysical studies on the various kinds of *Janus* PNAs, non-Janus PNAs and modified PNA oligomers:

- The *aeg* backbone in which the nucleobases attached at C_α through triazole linker effectively form sequence specific duplexes with DNA, similar to standard and well explored *aeg*-PNA in which the nucleobases are linked to backbone via tertiary amide link.

- The thermal stability (UV-T_m) of individual duplexes from amide and triazole faces were different, with triazole face duplex having higher stability than amide face duplex. In double duplexes from *Janus* PNA, T_m values of each face duplex was significantly enhanced. This pointed to the synergistic effect on the stability of duplexes when formed simultaneously. *The formation of double duplexes clearly provides the proof of concept for the hypothesis of design of Janus PNAs in the thesis.*

- The amide face *Janus* PNA with complementary DNA showed less T_m (amide face < trizolyl face). This is perhaps due to triazole ring aligning coplanar with the nucleobases and flexible to precisely orient the base pairs in duplex, indirectly strengthening the nucleobase to pair with complementary sequence. *The concomitant stacking of triazole rings during base pair formation may be the driving force for the enhanced thermal stability for duplex formation from triazole face.*

- Characteristic CD bands were seen for individual duplex formation from amide / triazole faces and simultaneous duplex formation from both faces in presence of appropriate complementary DNAs.

- The final CD profile in sequential titration experiments of *homo Janus* PNAs (T$_7$*jp-tz-C$_5$*) with complementary DNAs, DNA **1** (dA$_8$) and DNA **2** (dG$_6$) in both ways (first dA$_8$ addition followed by dG$_6$ and vice versa) are identical indicating the formation of same conformation of double duplex, irrespective of the order duplex assembly.

- The formation of complex is also seen in ESI MS (triplex, duplex and double duplex) as well as stoichiometry by UV Job plot experiments.

- ITC results of *homo Janus* PNA (T₇*jp-tz-C₅*) show the formation of triplex from amide face, duplex from triazolyl face, and duplex of triplex in presence of complementary DNAs (dA₈ and dG₆). In spite of higher UV-T_m of triazole duplex, the dissociation constant K_D values suggest that in *homo Janus* PNA duplexes, the triazole face duplex dissociates 6-10 times faster than the amide face duplex. The stoichiometry of binding (N) computed from ITC data in terms of number of base pairs nicely corresponded to the composition of envisaged duplexes. The enthalpy ΔH of duplexes from each face in individual *Janus* PNA complexes are in the range -20.0 to -40.0 kcal/mole. The range of free energy (ΔG) changes for different complexes was in a narrow range of -6 to -9 kcal/mole. The enthalpic contribution to free energy was slightly higher than entropic contribution in almost all complexes, suggesting that the *Janus* PNA:DNA binding reactions are perhaps driven by enthalpic considerations.

4.7 Summary

To summarize, this chapter deals with the biophysical investigation of *Janus* PNA monomers and various types of *Janus* PNA oligomers using Job plot, temperature dependent UV-visible spectroscopy, CD spectroscopy, ESI-MS and ITC. The next chapter deals with the i-motif and g-quadruplex formation abilities of these *Janus* PNA oligomers.
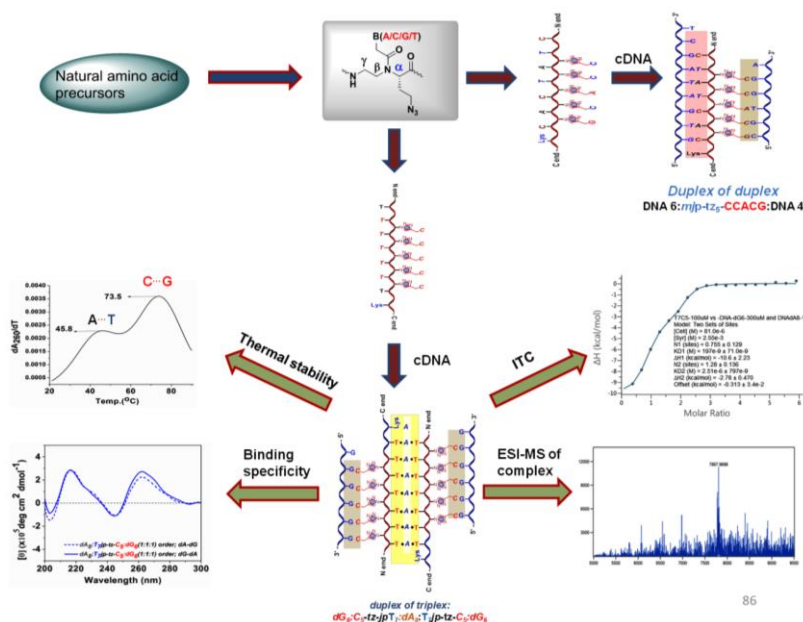
Figure 4.44 Summary of biophysical studies of *Janus* PNA oligomers

4.8 Experimental procedures

Chemicals: The unmodified and modified PNA oligomers were synthesized manually by general solid-phase PNA synthesis protocol using *Boc*-strategy as described in Chapter 2. Complementary DNA oligonucleotides were obtained commercially from Integrated DNA Technologies (IDT). Salts and reagents used in buffer preparation such as NaCl, Sodium cacodylate etc. were obtained from Sigma-Aldrich. The pH of the buffer solutions was adjusted using NaOH or HCl, from Sigma Aldrich.

4.8.1 UV-T_m measurement

UV-melting experiments were carried out on Varian Cary 300 UV spectrophotometer equipped with a peltier. The samples for T_m measurement were prepared by mixing the calculated amounts of respective oligonucleotides in the stoichiometric ratio (1:1, duplex) in sodium Cacodylate buffer (10 mM) and NaCl (10 mM), pH 7.2 to achieve a final strand concentration of 2 µM / 3 µM for each strand. The samples were annealed by heating at 90 °C for 10 min. followed by slow cooling to room temperature for at least 8-10 h and then refrigerated for at least 12 to 24 h. The samples (500 µL) were transferred to quartz cell and equilibrated at the starting temperature for 5 min. The OD at 260 nm was recorded in steps from 20-92 °C with temperature increment of 0.5 ºC /min. Each melting experiment was repeated at least twice. The normalized absorbance at 260 nm was plotted as a function of the temperature. The T_m was determined from the first derivative of normalized absorbance with respect to temperature and is accurate to ± 1.0 °C. The data were processed using Microcal Origin 8.5 fitted by sigmoidal curve using Boltzmann function for one face and biphasic dose response for two face binding. [The concentration of all oligonucleotides were calculated on the basis of absorbance from the molar extinction coefficients of the corresponding nucleobases i.e. T = 8.8 cm^2/µmol; C = 6.6 cm^2/µmol; G = 11.7 cm^2/µmol and A = 13.7 cm^2/µmol][30]

4.8.2 Circular Dichroism

CD spectra were recorded on JASCO J-815 spectropolarimeter connected with a peltier. The calculated amounts of PNA/*Janus* PNA oligomers and the complementary DNA were mixed together in stoichiometric ratio (1:1 for duplex) in sodium cacodylate buffer (10 mM) containing and NaCl (10 mM); pH 7.2 to achieve a final strand concentration of 5 or 10 µM for

each strand. The samples were annealed by heating at 90 °C for 10 min. followed by slow cooling to room temperature for at least 8-10 h. The cooled samples were transferred to refrigerator for at least 8 to 12 h. To record the CD spectra of PNA:DNA duplexes and single stranded PNAs, the temperature of circulating water was kept at 10 °C. The CD spectra were recorded as an accumulation of 3 scans from 300 to 190 nm using 2 mm cell, a resolution of 0.1 nm, band-width of 1 nm, sensitivity of 2 m deg, response of 2 sec and a scan speed of 100 nm/min.

4.8.3 Electrospray Ionisation Mass Spectrometer (ESI MS)

ESI-MS experiment was performed with Agilent 6540 QTOF MS instrument under ESI positive ion mode within mass range of 200-3200 m/z. Scan rate was kept 1Hz under fragmentor voltage 200 and capillary voltage 2500 V. Gas temp and gas flow were kept at 325 °C and 8 L/min respectively. Data acquisition and analysis software were done by Agilent Mass Hunter Version B.05.01 and Agilent Bioconfirm with Maximum Entropy Algorithm, 5000-10000 Da mass range and 1 Da mass step. Signal to noise ratio was 10; Minimum consecutive charge was 3 and baseline subtraction factor was kept 7.

Mass Spectrometry: Model Agilent 6540 UHD QTOF MS; Source ESI; Gas Temp (°C) 325; Gas Flow (l/min) 8; Nebulizer (psig) 30; Sheath Gas Temp (°C) 300; VCap (V) 2500; Nozzle Voltage (V) 0; Fragmentor (V) 200; Skimmer1: 45; Mass range (m/z) 100-3200; Scan speed (Hz) 1; Data type Profile; Software Masshunter workstation software v. B.05.01

Chromatography: Model Agilent Binary LC 1260; Column No column –Directly injected through UPLC; Mobile phase: A) Water (0.1% Formic Acid); B) Acetonitrile (LC-MS grade, J.T. Baker); Program type Isocratic 50% B, 0.2 mL/min; Injection volume 20 μL; Injection wash with methanol.

4.8.4 Isothermal Titration Calorimetry (ITC)

Thermodynamic properties such as enthalpy, entropy and binding constants of *Janus* PNA complexation to cDNA were determined using isothermal titration calorimetry (ITC). The binding interactions between *Janus* PNAs and their complementary DNA oligonucleotides were carried out on Malvern MicroCal PEAQ ITC instrument. All titration experiments were performed at 15 °C in 10 mM sodium cacodylate buffer (pH 7.2) containing NaCl (10 mM)

using same buffer solution to prepare all the solutions used in the experiment. The sample cell was loaded with *Janus* PNA solution and the reference cell contained only the buffer.

For determination the binding stoichiometry ratio for *homo Janus* PNA. each experiment performed at constant concentration (81 μM), the syringe was loaded with DNA solution (40 μL) 220 μM to 1.5 mM depend upon experiment (concentration is given in terms of nucleobase for comparison) depending upon experiments. The instrument was equilibrated at 15 °C until the baseline was flat and stable. *Janus* PNA was taken in the cell for binding experiment with cDNA to each face and total of 19 injections (2 μL/injection) were performed into a solution of *Janus* PNA in the cell.

N obtained from the (No. of nucleobases involved in binding from DNA / No. of nucleobases involved in binding from *Janus* PNA).

4.9 References

1. B. Ranjbar, P. Gill, *Chem. Biol. Drug Des.* **2009**, *74*, 101-120.

2. Nakano, S.-i.; Kanzaki, T.; Sugimoto, N. *J. Am. Chem. Soc.* **2004**, *126*, 1088-1095.

3. Leavitt, S.; Freire, E. *Curr. Opin. Struct. Biol.* **2001**, *11*, 560-566.

4. Ren, C.; Bailey, A. O.; VanderPorten, E.; Oh A.; Phung, W.; Mulvihill, M. M.; Harris, S. F.; Liu, Y.; Han, G.; Sandoval, W. *Anal. Chem.* **2019**, *91*, 903-911.

5. Loo, R. R. O.; Udseth, H. R.; Smith, R. D. *J. Phys. Chem.* **1991**, *95*, 6412-6415.

6. Pan, J.; Xu, K.; Yang, X.; Choy, W.-Y.; Konermann, L. *Anal. Chem.* **2009**, *81*, 5008-5015.

7. Benesch, J. L. P.; Ruotolo, B. T. Simmons, D. A.; Robinson, C. V. *Chem. Rev.* **2007**, *107*, 3544-3567.

8. Micsonai, A.; Wien, F.; Kernya, L.; Lee, Y.-H.; Goto, Y.; Refregiers, M.; Kardos, J. *Proc. Natl. Acad. Sci.U.S.A.* **2015**, *112*, E3095-E3103.

9. Kantonen, S. A.; Henriksen, N. M.; Gilson, M. K. *Biochim. Biophys. Acta, Gen. Subj.* **2017**, *1861*, 485-498.

10. Sturtevant, J. M.; *Proc. Natl. Acad. Sci. U. S. A.* **1977**, *74*, 2236-2240.

11. Huang, C. Y. *Methods Enzymol. 87* (Ed.: Purich D. L.), Academic Press, **1982**, 509-525.

15. Gasser, G.; Husken, N.; Koster, S. D.; Metzler-Nolte, N. *Chem. Commun.* **2008**, 3675-3677;

16 Kuijpers, B. H. M.; Groothuys, S.; Hawner, C.; Dam, J. t.; Quaedflieg, P. J. L. M.; Schoemaker, H. E.; Delft, F. L. v.; Rutjes, F. P. J. T. *Org. Process Res. Dev.* **2008**, *12*, 503-511.

17. (a) Tornøe, C. W.; Christensen, C.; Meldal, M. *J. Org. Chem.* **2002**, *67*, 3057-3064; (b) Chow, H. Y.; Zhang, Y.; Matheson, E.; Li, X. *Chem. Rev.* **2019**.

18. Kim, S. K.; Nielsen, P. E.; Egholm, M.; Buchardt, O.; Berg, R. H.; Norden, B. *J. Am. Chem. Soc.* **1993**, *115*, 6477-6481.

19. Mitra, R.; Ganesh, K. N. *J. Org. Chem.* **2012**, *77*, 5696-5704.

20. Dragulescu-Andrasi, A.; Rapireddy, S.; Frezza, B. M.; Gayathri, C.; Gil, R. R.; Ly, D. H. *J. Am. Chem. Soc.* **2006**, *128*, 10258-10267.

21. Gangamani, B. P.; Kumar, V. A.; Ganesh, K. N. *Tetrahedron* **1996**, *52*, 15017-15030.

22. Krupnik, O. V.; Guscho, Y. A.; Sluchanko, K. A.; Nielsen, P. E.; Lazurkin, Y. S. *J. of Biomol. Struct. Dyn.* **2001**, *19*, 535-542.

23. Schwarz, F. P.; Robinson, S.; Butler, J. M. *Nucleic Acids Res.* **1999**, *27*, 4792-4800;

24. Ratilainen, T.; Holmén, A.; Tuite, E.; Nielsen, P. E.; Norden, B. *Biochemistry* **2000**, *39*, 7781-7791.

25. Sato, T. Sakamoto, N.; Nishizawa, S. *Org. Biomol. Chem.* **2018**, *16*, 1178-1187;

26. Gupta, P.; Muse, O.; Rozners, E. *Biochemistry* **2012**, *51*, 63-73;

27. (a) Lomzov, A. A.; Vorobjev, Y. N.; Pyshnyi, D. V. *J. Phys. Chem. B* **2015**, *119*, 15221-15234; (b) Sen, A.; Nielsen, P. E. *Biophys J.* **2006**, *90*, 1329-1337.

28. Bartold, K.; Pietrzyk-Le, A.; Golebiewska, K.; Lisowski, W.; Cauteruccio, S.; Licandro, E.; D Souza, F.; Kutner, W. *ACS Appl. Mater. Interfaces* **2018**, *10*, 27562-27569.

29. Tomac, S.; Sarkar, M.; Ratilainen, T.; Wittung, P.; Nielsen, P. E.; B. Norden, Graslund, A. *J. Am. Chem. Soc.* **1996**, *118*, 5544-5552.

30. Jain, D. R.; Ganesh, K. N. *J. Org. Chem.* **2014**, *79*, 6708-6714.

4.10 Appendix III

4.10.1 CD spectra single stranded of DNA, PNA and *Janus* PNA

The *homo Janus* PNA $T_7jp\text{-}tz\text{-}C_5$ (**JP 1**) and control PNAs - *aeg* PNA-T_7 (PNA **15**), *aeg-tz* , $p7\text{-}tz\text{-}C_5$ (PNA **13**), oligomers exhibited very low induced CD signals although JP **1** and PNA **13** have chiral centers, but removed far away from the nucleobase chromophore
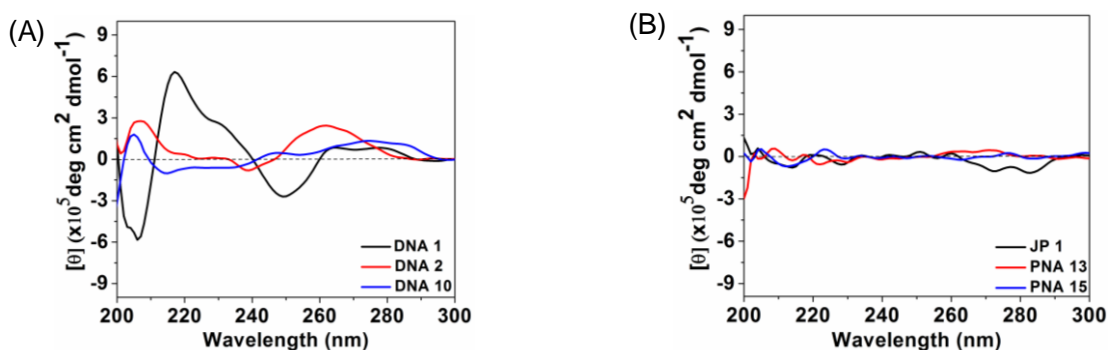
Figure 4.45 CD spectra of (**A**) *ss*DNAs: dA_8 (DNA **1**), dG_6 (DNA **2**) and dC_6 (DNA **10**) and (**B**) *ss*PNAs: PNA T_7 (PNA **15**), $p7\text{-}tz\text{-}C_5$ (PNA **13**), *homo Janus* PNAs: $T_7jp\text{-}tz\text{-}C_5$ (**JP 1**); Buffer: 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM)
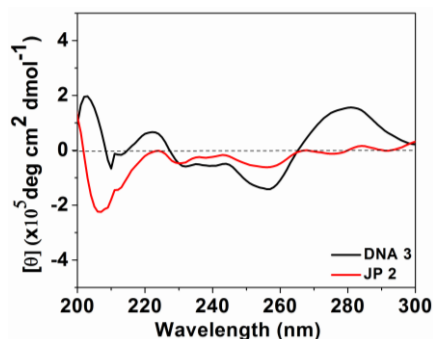
Figure 4.46 CD spectra of ssDNA DNA **3** (5'-GTGATCT-3') and *chimeric Janus* PNA: $mjp\text{-}tz\text{-}C_5$ (**JP 2**); Buffer: 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM.
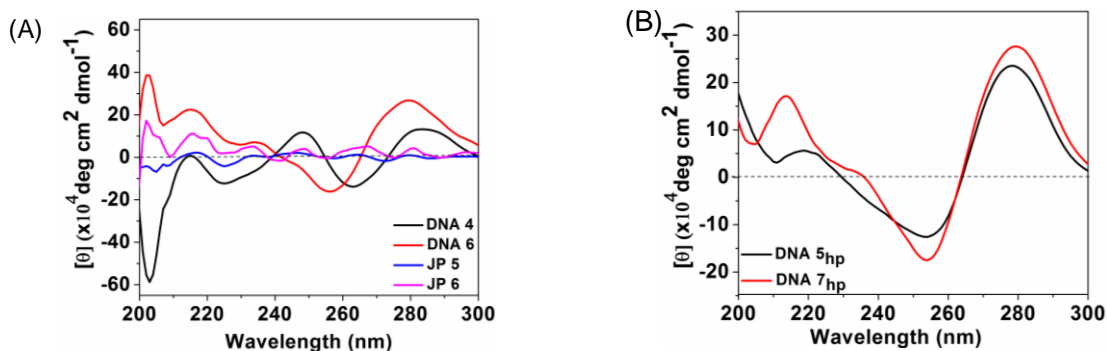
Figure 4.47 CD spectra of **(A)** single stranded DNA **4,** DNA **6,** m*jp-tz₄-CACG* (**JP 5**) *and* m*jp-tz₅-CCACG* (**JP 6**) **(B)** hairpin DNA, DNA **5**ₕₚ, DNA **7**ₕₚ; Buffer: 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM)

4.10.2 ITC dilution experiments for JP 1 or T₇jp-*tz*-C₅ *(DNA vs buffer titration)*
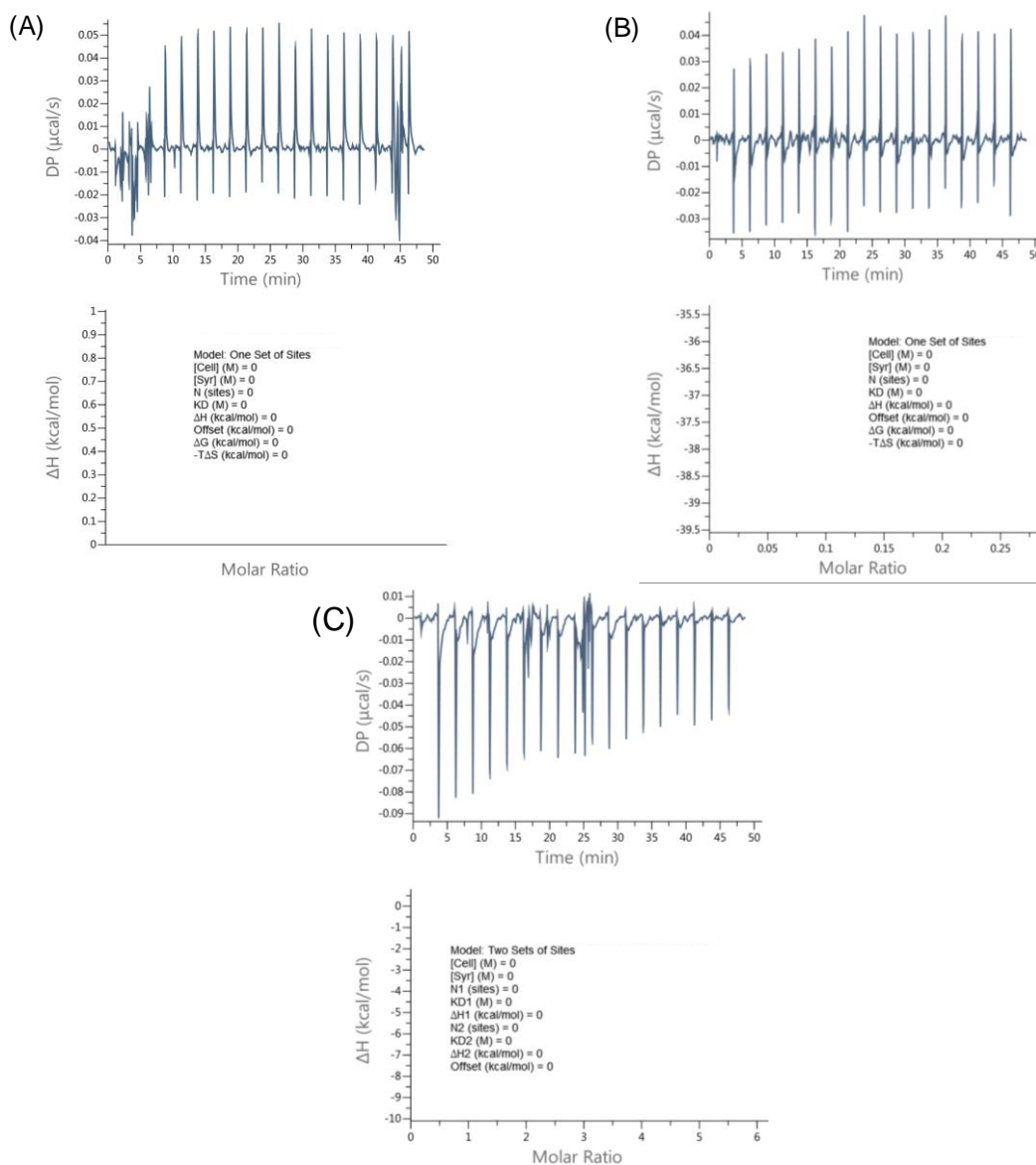
Figure 4.48 ITC dilution experiments of DNA vs buffer at 15 °C (A) DNA 1 (1.05 mM) vs buffer **(B)** DNA 2 (0.75 mM) vs buffer **(C)** mixture of DNA **1** (1.05mM) + DNA **2** (1.5mM) (1:2) vs buffer and titration.

4.10.3 ITC dilution experiments for m*jp-tz₄-CACG* or JP 5 *(DNA vs buffer titration)*
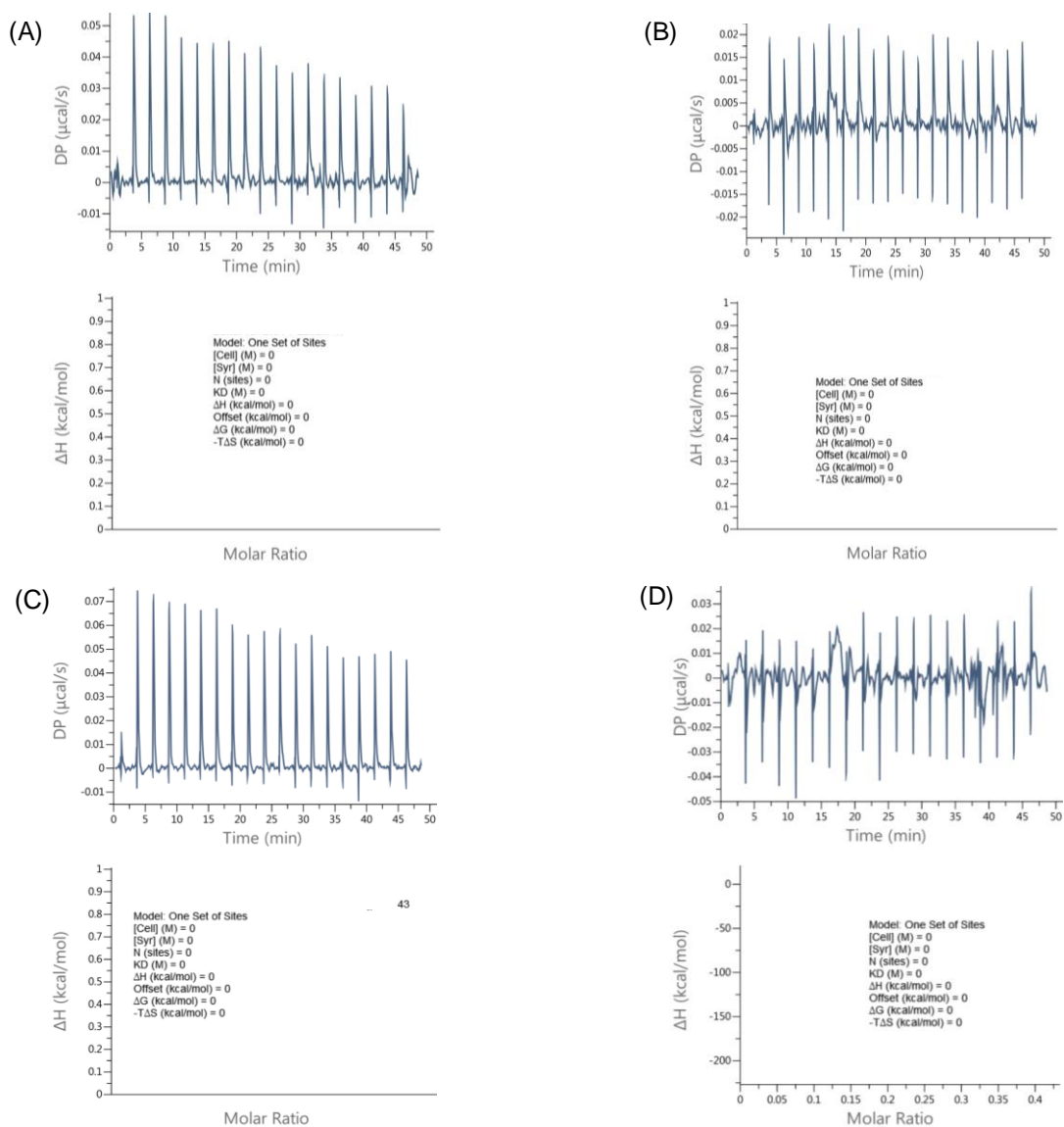
Figure 4.49 ITC dilution experiments at 15 ᵒC (A) DNA **3** (1.10 mM) vs buffer **(B)** DNA **4** (0.88 mM) vs buffer **(C)** mixture of [DNA **3** (1.1mM) + DNA **4** (0.88 mM) (1:1) vs buffer **(D)** DNA **5**ₕₚ (1.98 mM) vs buffer titration.

4.10.4 ITC dilution experiments for m*jp-tz₅-CCACG* or JP 6 *(DNA vs buffer titration)*
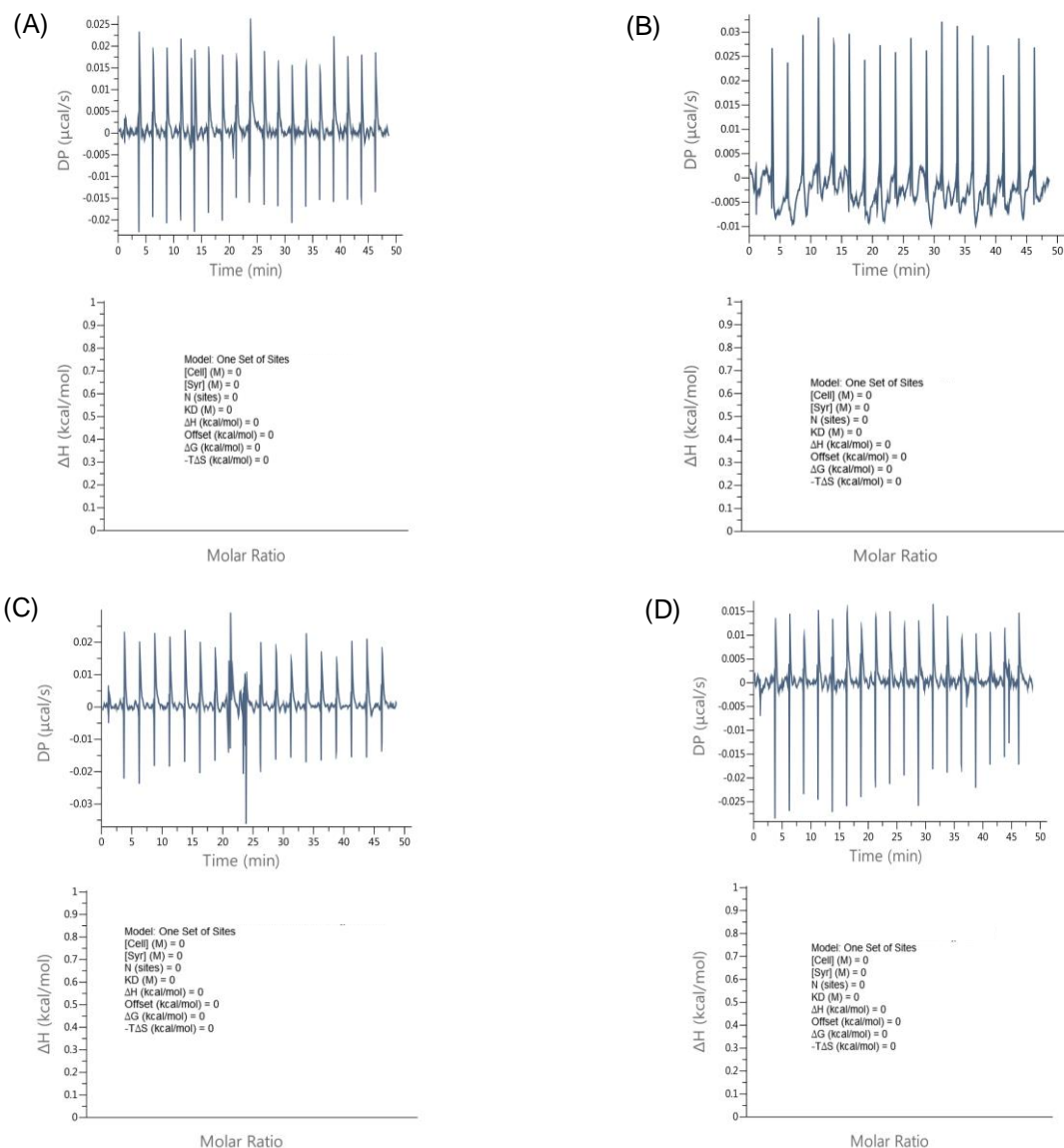
Figure 4.50 ITC dilution experiments at 15 °C (A) DNA **6** (280 μM) vs buffer **(B)** DNA **4** (200 μM) vs buffer **(C)** mixture of [DNA **6** (0.21 mM) + DNA **4** (0.15 mM) (1:1) vs Buffer **(D)** DNA **7**$_{hp}$ (1.20 mM) vs Buffer titration.

Chapter 5

i-Motif and G-Quadruplex Structures from *Janus* PNA Oligomers and their Morphological study by FESEM

5.0 Introduction

DNA, a carrier of genetic information, is a right-handed double helical structure comprised of Watson-Crick base pairing: adenine (A) with thymine (T) and cytosine (C) with guanine (G), under ordinary physiological conditions.[1] Shortly after the discovery of DNA double helix, non-canonical DNA structures consisting of non-Watson-Crick base pairing were evident in repetitive DNA sequences.[2] DNA has a well-known propensity to adopt alternative non-B-form conformations in vitro, including interdigitated motif (i-motif) structures[3] G-quadruplex (G4)[3] and supramolecular self-assembly. Recently several nanodevices have been built for electronics, biosensors, and biomimetic applications including their applications in supramolecular chemistry.[4]
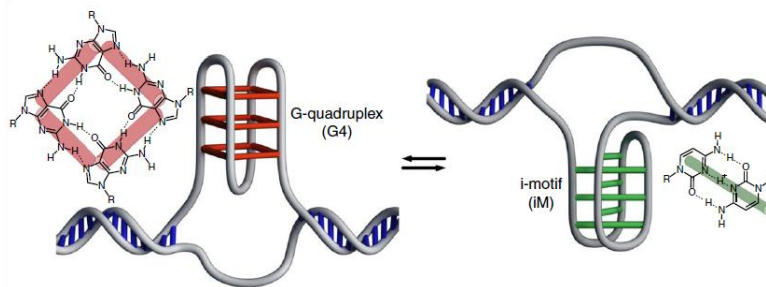
Figure 5.1 Schematic G4 and i-Motif structure[5]

5.1 i-Motif

DNA exists as a double-stranded helix in the cell,[1] sometimes the cytosine-rich strands complementary to the guanine-rich strand is known to fold into i-motif structures under certain conditions.[6] The discovery of i-motif with a detailed NMR structure was reported by Gehring *et al* in 1993.[7] The structure of i-motif consists of two parallel duplexes held together by hemi-protonated cytosine pairs (C:CH$^+$) (Figure 5.2).[6,7] These C:CH$^+$ pairs interdigitate into each other with opposite orientation and in C:CH$^+$ pairs cytosine is hemi-protonated at the N3 position.

Until now, i-motif structures are the only known nucleic acid structures stabilized by base intercalation.[8] The structure has two broad (~1.5 nm) and two narrow grooves (~0.7 nm). Depending on the number of DNA strands involved, it has been demonstrated that the DNA sequences with consecutive runs of cytosines can fold into intramolecular or intermolecular i-motif at acidic pH.7.[9] A DNA strand with one, two or four cytosine repeats can fold into tetrameric, dimeric, and monomeric (or intramolecular) i-motif structures, respectively.[6] i-Motif

structures also demonstrate structural polymorphism in terms of loop length and number of intercalated base pairs in different gene promoters.[10] This structural polymorphism may offer an important opportunity for small molecule targeting.[10] Besides the fully folded i-motif structures, a recent study reported the formation of intramolecular structure in less than three tandem C-rich repeats highlighting the potential competing structures.[11]
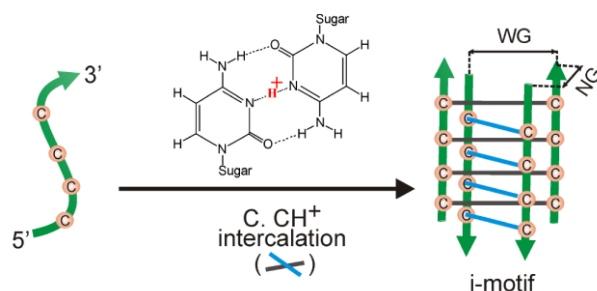
Figure 5.2 Formation of an i-motif structure by intercalation of C.CH[+] pairs. 'WG' and 'NG' depicts wide and narrow grooves respectively[11]

It has also been shown that PNA alone[12] or in association with DNA, can form a hybrid i-motif, and has very different properties compared to its DNA analog.[11] Although the formation of i-motif structure is favored at acidic pH, its formation has been reported at neutral and slightly basic pH under molecular crowding conditions.[13] PNA C–C[+] tetraplex properties of unmodified PNA sequences TC$_4$ and TC$_8$ are observed analogous to the iso-sequential DNA, but with higher thermal stability in the acidic pH range.[12]

Based on these observations one can design and utilize *Janus* PNA that can interestingly play two roles simultaneously from each face of backbone, forming i-motif from one face (either amide-face or modified-triazole) and duplex from other face in presence of cDNA/RNA/*aeg* PNA or *Janus* PNA and such complexes can be termed duplex fused tetraplex (Figure 5.3).

Chimeric JP   *duplex fused tetraplex*
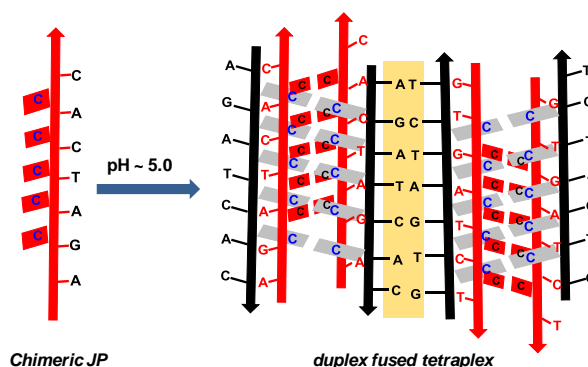
Figure 5.3 Duplex fused tetraplex possible from *Janus* PNA. *Janus* PNAs with homo cytidines on one face and hetero base sequence on another face.

5.2 G-quadruplex

It is known that many promoter regions of DNA contain repeating sequences capable of forming structures other than double helix. The DNA sequences with contiguous guanine repeats have the potential to form G-quadruplex.[3] The formation of G-quartet, a building block of G quadruplex structure, was first identified in 1962 on the basis of the aggregation of 5' monophosphate guanosines to form a guanosine gel.[14] When two or more G-quartets are stacked on top of each other, a tetrameric G-quadruplex is formed (Figure 5.4). The structure is stabilized by monovalent metal cation, typically K^+ or Na^+, by coordinating with O^6 carbonyl of the guanines.[15] G-quadruplex structures can assume uni-, bi-, or tetramolecular types.[16] Formation of a unimolecular structure requires a sequence with four G-rich stretches separated by at least one nucleotide. Upon folding into a G-quadruplex structure, the nucleotides between two tandem G-rich repeats form loops. Based on the strand orientation, a G-quadruplex can be parallel (the same polarity for all strands) or antiparallel (each strand has opposite polarity with respect to the two adjacent strands). (Figure 5.4) highlights general types of G-quadruplexes based on strand orientation. To date, different G-quadruplex structures have been reported in the human genome. G-quadruplex structures are highly polymorphic. A specific conformation depends on the DNA sequence, strand orientation, loop size, and solution factors such as cation species and molecular crowding conditions.[17]
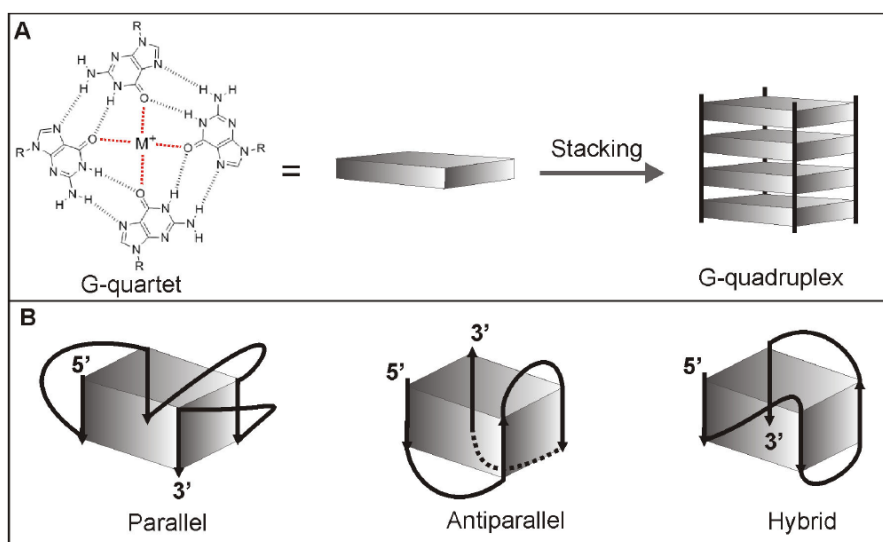
Figure 5.4 (A) Formation of G-quadruplex by stacking of G-quartets. **(B)** Three general types of G-quadruplexes based on DNA strand orientation.[17]

Peptide nucleic acid (PNA)[18] with *aeg* backbone are interesting to understand the role of the backbone in inducing the self-assemblies. The G-quadruplex derived from aminoethyl glycine *aeg*-PNA is found to be more stable than that of DNA.[19] Based on literature survey and duplex forming properties of *Janus* PNAs described in earlier chapters, we were motivated for synthesis of *Janus* PNA that can interestingly play two simultaneous roles based on nucleobases from each face of backbone.

(i)   G-quadruplex from one face (either amide-face or modified-triazole) and duplex from other face in presence of cDNA/RNA or PNA, leading to duplex fused tetraplexes (Figure 5.5A).

(ii)   Formation of G quadruplex from both face (amide-face as well as modified-triazole) simultaneously resulting in backbone fused G-tetraplexes (Figure 5.5B).
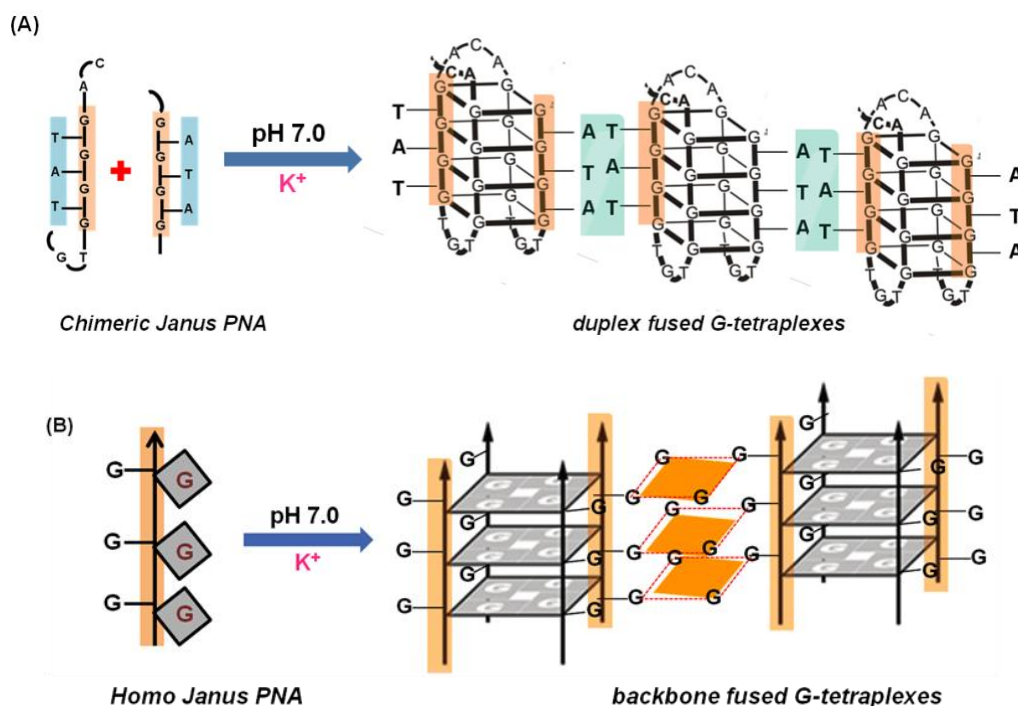
Figure 5.5 *Janus* G-quadruplex (**A**) Duplex fused G-tetraplexes (**B**) Backbone fused G-tetraplexes

The preceding chapter discussed the study of rationally designed C_α-substituted triazolyl derivatives *Janus* PNA analogs such as *homo Janus* PNAs, *chimeric hetero Janus* PNAs and *self-complementary hetero Janus* PNAs for their duplex forming abilities. This chapter reports studies on their ability to form i-motif and G-quadruplex structures in comparison with control

aeg PNA/*aeg-tz*-PNA/DNA. The last section discusses their self assembling properties through their morphological features probed using microscopic techniques such as FESEM and TEM.

5.3 Biophysical techniques used to study self-assembly of *Janus* PNA-G/C oligomers

Spectroscopic techniques such as UV spectroscopy for pK_a determination, temperature dependent UV absorbance (melting) to determine stability of i-motif and G-quadruplex, CD spectrometry to characterize i-motifs and G quadruplex structures are useful to determine the secondary structures of nucleic acids. Field emission scanning electron microscopy (FESEM) is used for characterizing morphological structures induced by supramolecular assembly of *Janus* PNA:DNA complexes.

5.4 Objectives of the present work

Based on the properties and special attributes of *Janus* PNAs described earlier, the objectives of this Chapter is to explore homo-G_n and homo-C_n oligomeric *Janus* PNA sequences (Table 5.1) and their variations to evaluate their potential for formation of fused duplex and tetraplex structures with cDNA sequences using various biophysical techniques. The specific objectives of this section are:

➢ Determination of the pK_a of N3 of cytosines in *homo C*-oligomers *p*7-*tz*-C_5 PNA **13** and PNA-C_5 (PNA **16**) and dC_6 (DNA **10**) by pH-dependent UV spectroscopy

➢ Studying the thermal stability of i-motif of T_7*jp-tz*-C_5 (**JP 1**), *p*7-*tz*-C_5 (PNA **13**), PNA-C_5 (PNA **16**) and dC_6 (DNA **10**) by pH-dependent UV-T_m

➢ Investigation of salt-dependent formation and stability of G-quadruplex of m*jp-tz*-G_4 (**JP 4**), *p*7-*tz*-G_5 (PNA **14**) and PNA-G_5 (PNA **17**) by salt-dependent UV-T_m

➢ Temperature-dependent CD experiment of *p*7-*tz*-G_5 (PNA **14**) to investigate the formation of G-quadruplex

➢ Supramolecular assembly studies of *homo Janus* PNA T_7*jp-tz*-C_5 (**JP 1**), *chimeric Janus* PNAs m*jp-tz*-C_5 (**JP 2**) and m*jp-tz*-G_4 (**JP 4**), *self-complementary Janus* SCM-PNA (**JP 7**), *p*7-*tz*-C_5 (PNA **13**), amino PNA *p*7-*tz*-G_5 (PNA **14**), *aeg*-PNA-C_5 (**PNA 16**), *aeg*-PNA-G_5 (PNA **17**) (Table 5.1) and cDNA (Table 5.2) in water.

Table 5.1 *Janus* PNAs / *aeg-tz* PNA and *aeg* PNA used for study of i-motif, G-quadruplex, and self-assembly

Entry	Sequence code	*Janus* / *aeg-tz* /*aeg* PNA sequences
1	**T$_7$jp-tz-C$_5$ (JP 1)**	
2	**mjp-tz-C$_5$ (JP 2)**	
3	**mjp-tz-G$_4$ (JP 4)**	
4	**SCM-JP (JP 7)**	
5	**p7-tz-C$_5$ (PNA 13)**	
6	**p7-tz-G$_5$ (PNA 14)**	
7	**PNA-C$_5$ (PNA 16)**	
8	**PNA-G$_5$ (PNA 17)**	

m = mix nucleobases, *jp* = *Janus* PNA, *tz* = triazolyl, *SCM-JP* = *Self Complementary hetero Janus* PNA, A = Adenine, C = Cytosine, G = Guanine, T = Thymine

Table 5.2 Complementary DNA used for i-motif, G-quadruplex and self assembly

Entry	DNA	Sequence (5' to 3')	Type
1	DNA **1**	AAA AAA AA or (**dA$_8$**)	Anti-parallel
2	DNA **2**	GGG GGG or (**dG$_6$**)	Anti-parallel
3	DNA **3**	GTG ATC T (**cDNA**)	Anti-parallel
4	DNA **8**	GAG GAG AC	Anti-parallel
5	DNA **9**	TCC TCT	Anti-parallel
6	DNA **10**	CCC CCC (**dC$_6$**)	Anti-parallel

5.5 i-Motif and G-quadruplex study of *Janus* PNAs

This section describes the studies of pK$_a$ determination and i-motif formation of homo C-oligomers: *p7-tz-C$_5$* (PNA **13**)**,** PNA-C$_5$ (PNA **16**) and dC$_6$ (DNA **10**) and and G-quadruplex formation in G-rich oligomers m*jp-tz-G$_4$* (**JP 4**), PNA **14** (*p7-tz-G$_5$*), *aeg*-PNA-G$_5$ (PNA **17**).

5.5.1 pK$_a$ determination of C-oligomers

The ability of PNA and DNA C-oligomers to form i-motif depends on formation of C-CH$^+$ homo base pairs between free C and protonated C (N3-H$^+$, C$^+$). It is therefore important to determine the pK$_a$ of N3 in C in different PNA oligomers and its variation as a function of backbone structure. There are several methods for the determination of pK$_a$ like traditional potentiometry, UV-Vis spectrometry *etc*. In the present studies, the UV-Vis absorption method has been used, since protonation of C changes its UV absorption properties. In order to determine aqueous pK$_a$ values of C in C$_n$–oligomers DNA **10**, PNA **13** and PNA **16** at concentration of 10 μM, the ratio of UV absorbance of free C (271 nm) and protonated C (279 nm) (Figure 5.6A) was plotted as a function of pH (Figure 5.6B). The sigmoidal ratio plot of absorbance (A$_{271}$/A$_{279}$) vs pH were fitted (Boltzmann) using Origin 8.5 software, and from the maxima in the derivative of the fitted curve the pK$_a$ value (5.6C), was obtained for each oligomer (Table 5.3).
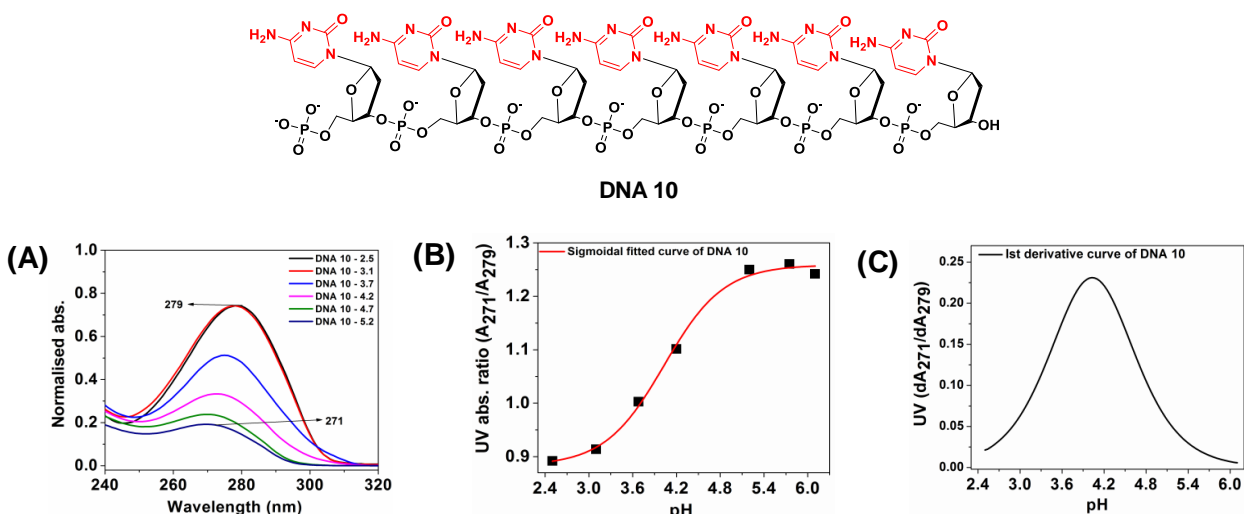

DNA 10

Figure 5.6 pK$_a$ determination of C (N3) in DNA **10** (dC$_6$) (A) pH dependent UV spectra; (B) A$_{271}$/A$_{279}$ sigmoidal fitted vs pH (C) Ist derivative (A$_{271}$/A$_{279}$) vs pH.
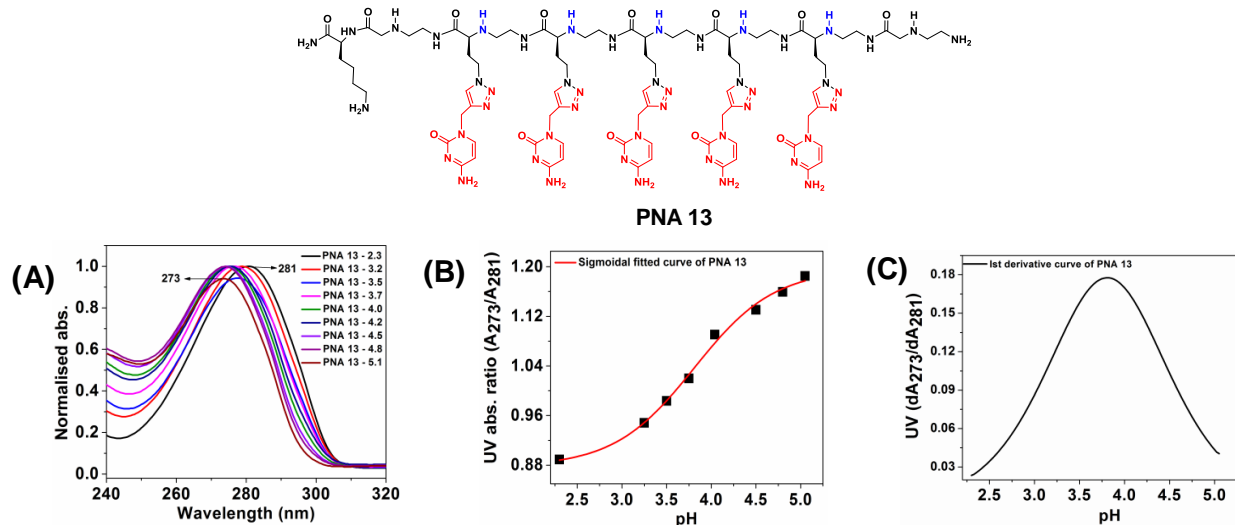
PNA 13

Figure 5.7 pKa determination of C (N3) in *p*7-*tz*-*C₅* (*aeg-tz* PNA **13**) (A) pH dependent UV spectra (B) A_{273}/A_{281} ratio of abs.maxima vs pH (C) Ist derivative (A_{273}/A_{281}) vs pH.
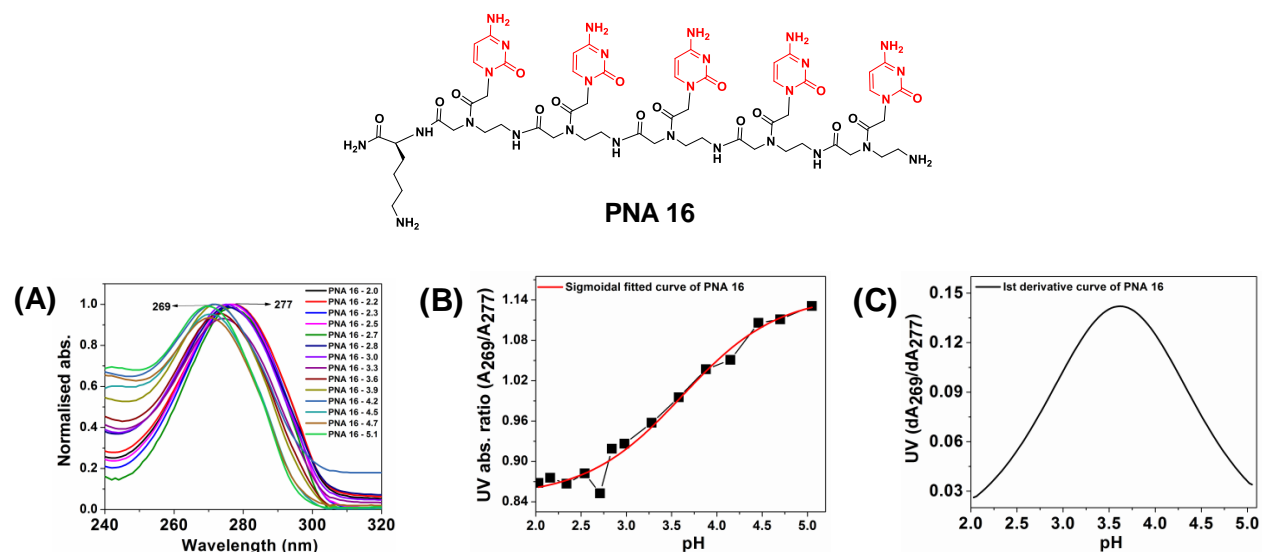
PNA 16

Figure 5.8 pKa determination of C (N3) in PNA-C₅ (PNA **16**) (A) pH-dependent UV-Vis spectra (B) A_{269}/A_{277} sigmoidal fitted vs pH (C) Ist derivative (A_{269}/A_{277}) vs pH.

The pKas obtained for C (N3) in various C-oligomers are shown in Table-3. It is seen that the pka of N3 of cytosine in PNA oligomers **13** and **16** is lower by approximately 0.4 units compared to that of C in DNA **10**. Thus C (N3) in PNAs is more acidic than that in DNA. Since protonation of N3 is important for i-motif, the lower pKa of N3 of C in PNA leads to lower stability of PNA i-motif at pH 7.0 compared to that in DNA (less amount of protonated C in PNA at pH 7.0 compared to that in DNA).

Table 5.3 Summary of pk$_a$ (N3 of cytosine)

Entry	oligomers	pk$_a$
1	dC$_6$ (DNA **10**)	4.03
2	*p*7-*tz*-*C$_5$* (*aeg-tz* PNA **13**)	3.81
3	PNA-C$_5$ (*aeg*-PNA **16**)	3.60

5.5.2 i-Motif studies by pH-dependent UV-T_m

The UV-T_m studies of the *i*-motif formation by *homo Janus* PNA T$_\gamma$*jp-tz-C$_5$* (**JP 1**), *aeg-tz* PNA *p*7-*tz*-*C$_5$* (PNA **13**), *aeg*-PNA-C$_5$ (PNA **16**) and dC$_6$ (DNA **10**) were done at acidic pH [sodium acetate (100 mM) buffer (pH 3.0 - 5.0)] by monitoring absorbance at 295 nm which is characteristic of C$^+$ (N3-H$^+$).[12]

Figure 5.9 shows UV-temperature plot of *aeg*-PNA-C$_5$, dC$_6$, *p*7-*tz*-*C$_5$* and *Janus* PNA T$_\gamma$*jp-tz-C$_5$* monitored at 295 nm at acidic pH 3.27 and the corresponding T_m of the *i*-motif structures were evaluated from first derivative plots. dC$_6$ and *aeg*-PNA-C$_5$ showed broad transitions and negative inverse first derivative curves indicated the formation of *i*-motif tetraplexes with T_m 54.6 ^{o}C and 74.5 ^{o}C respectively and T_m of *aeg*-PNA-C$_5$ (PNA **16**) > T_m of dC$_6$ (DNA **10**). The *Janus* PNA T$_\gamma$*jp-tz-C$_5$* (**JP 1**) and amino-triazole *p*7-*tz*-*C$_5$* (PNA **13**), *aeg*-PNA-C$_5$ (PNA **16**) failed to show any transition in the temperature range 20 ^{o}C to 90 ^{o}C.
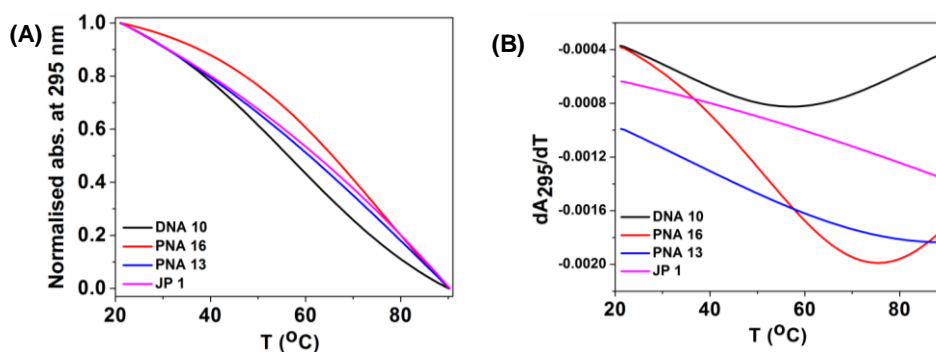

Figure 5.9 UV-T_m of C$_n$-tetraplexes of DNA/PNA/*aeg-tz* PNA and *Janus* PNAs at pH 3.27 **(A)** normalised absorbance at 295 nm **(B)** Ist derivative curve (dA$_{295}$/dT).

Figure 5.11, shows UV-temperature data for C-oligomers at pH 5.0, monitored at 295 nm. Inverse sigmoidal transition characteristic of i-motif were seen for all four C-oligomers. The order of T_m was *Janus* PNA T$_\gamma$*jp-tz-C$_5$* (57.8 ^{o}C) ~ *aeg-tz* PNA **13** (58.1 ^{o}C) > dC$_6$ (42.7 ^{o}C) with *aeg*-PNA-C$_6$ (PNA **16**) showing very broad and uncertain transition. From these experiments, it

is clearly seen that *Janus* PNA T_7jp-tz-C_5 (**JP 1**) and *aeg-tz* PNA **13** form *i*-motif with better stability compared to dC$_6$, while *aeg*-PNA **16** forms weak i-motif structures.
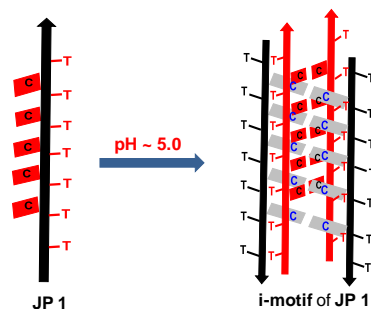

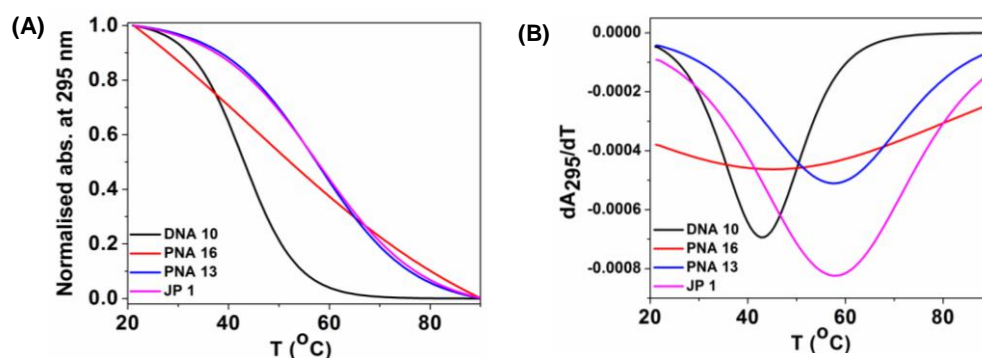**Figure 5.10** i-motif formation by **JP 1** (T_7jp-tz-C_5)

Figure 5.11 UV-T_m of C$_n$-tetraplexes of DNA/PNA/*aeg-tz* PNA and *Janus* PNAs at pH 5.0 **(A)** normalised absorbance at 295 nm **(B)** Ist derivative curve (dA$_{295}$/dT).

In contrast to results at pH 5.0, at physiological pH 7.5 (sodium cacodylate buffer 100 mM), the UV-Temperature plots for all oligomers were positive sigmoidal at 295 nm (compared to negative sigmoidal at lower acidic pH) (Figure 5.12), with very sharp transitions as seen in first derivative curves. The positive peak in first derivative curves corresponds to increase in absorbance at 295 nm with temperature as expected for duplex formation (opposite of that seen for *i*-motif formation), and the order of T_m stability is reverse of that seen earlier for *i*-motif structures: DNA **10** (43.5 °C) > *Janus* PNA **JP 1** (33.4 °C) > *aeg-tz* PNA **13** (32.2 °C) > *aeg*-PNA **16** (31.6 °C). Such duplex structures result from C-C$^+$ base pairing, but without interdigitation. Table 5.4 summarizes the results of all UV-T_m results for various DNA and PNA C-oligomers.
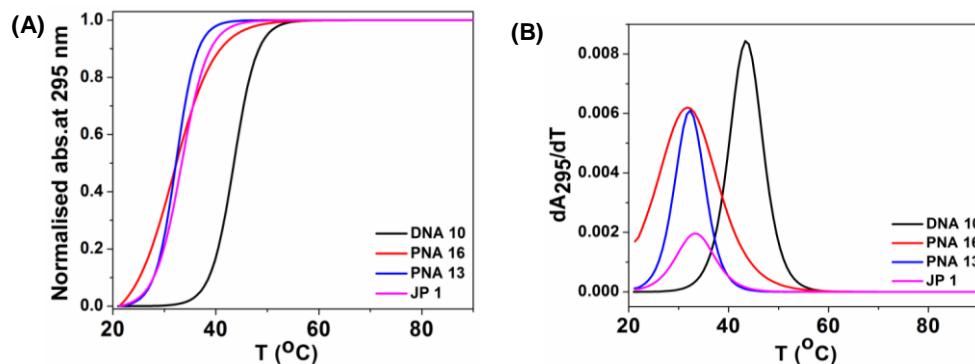
Figure 5.12 UV-T_m of DNA/PNA/ *aeg-tz* PNA and *Janus* PNAs at pH 7.36 **(A)** normalised absorbance at 295 nm **(B)** Ist derivative curve (dA$_{295}$/dT).

Table 5.4 Summary of pH-dependent UV-T_m (^{o}C) of *homo Janus* PNA *T_7jp-tz-C_5* (**JP 1**), *aeg-tz* PNA *p7-tz-C_5* (**PNA 13**) PNA **16** (PNA-C_5) and DNA **10** (dC$_6$) oligomers at 295 nm abs**.**

Entry	Oligomers	pH 3.27[#]	pH 5.0[#]	pH 7.34*
1	**dC$_6$**	54.6	42.7	43.5
2	**PNA C_5**	74.5	N.T.	31.6
3	*p7-tz-C_5*	N.T.	58.1	32.2
4	*T_7jp-tz-C_5*	N.T.	57.8	33.4

[#] i-Motif formation; *Duplex formation, N.T. = No Transition

5.5.3 UV-T_m studies of quadruplex formation by G$_n$-oligomers: *Janus* PNA m*jp-tz-G_4* (**JP 4**) *p7-tz-G_5* (**PNA 14**) and **PNA-G$_5$** (**PNA 17**)

This section discusses G-quadruplex formation by homo G-oligomers of *Janus* PNA m*jp-tz-G_4* (**JP 4**), *p7-tz-G_5* (**PNA 14**) and comparison with *aeg*-PNA G$_5$ (**PNA 17**). As pointed out earlier, G$_4$-tetraplexes are favored in presence of metal ions Na$^+$ or K$^+$, which co-coordinatively bind the O^6 of carbonyls of tetrameric G residues. Hence formation of G$_4$-tetraplexes from different G-oligomers *Janus* PNA m*jp-tz-G_4*, amino-triazole *p7-tz-G_5* (**PNA 14**) and *aeg*-PNA-G$_5$ (**PNA 17**) were examined by salt concentration-dependent UV-T_m, with each oligomer strand concentration of 5 μM at pH 7.1. The transitions were monitored at 295 nm which is characteristic band for G-tetraplex in 20 mM sodium cacodylate buffer having 50 mM and 100 mM KCl concentration (Figure 5.13 and Figure 5.14).
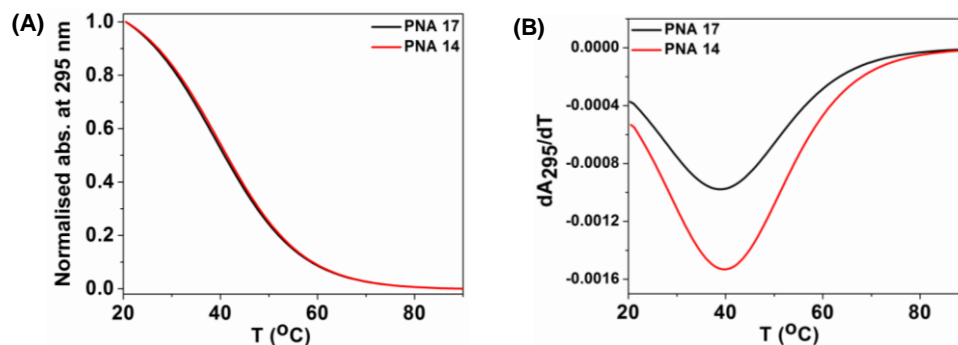
Figure 5.13 UV-T_{m} of PNA G-oligomers *p7-tz-G_5* (PNA **14**) and *aeg*-PNA **17** in 20 mM sodium cacodylate buffer having 50 mM KCl, pH 7.1 **(A)** normalised abs. at 295 nm **(B)** 1st derivative dA$_{295}$/dT curve.

Figure 5.14 UV-T_{m} of PNA G-oligomers *p7-tz-G_5* (PNA **14**) and aeg-PNA **17** in 20 mM sodium cacodylate buffer having 100 mM KCl, pH 7.1 **(A)** normalised abs. at 295 nm **(B)** 1st derivative dA$_{295}$/dT curve.
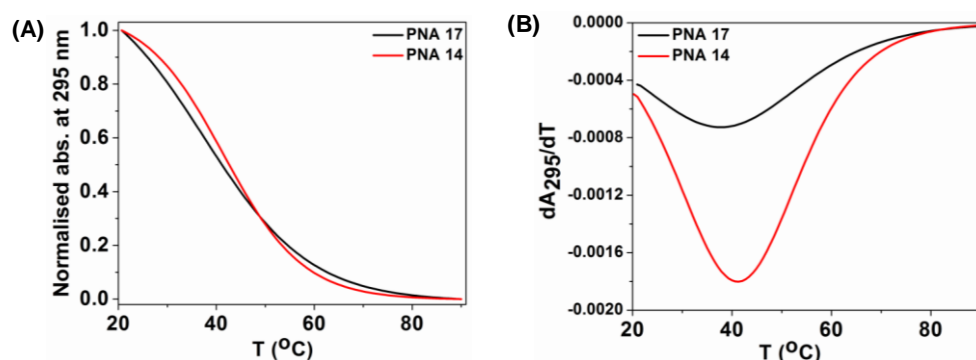
The UV-T_{m} experiments monitored at 295 nm[19] showed formation of G-tetraplexes from PNA G-oligomers *aeg* PNA **14** (*p7-tz-G_5*) and PNA-G$_5$ (PNA **17**) due to inverse sigmoidal transition and stability differences in presence of 50 mM and 100 mM KCl concentrations were negligible. The *Janus* PNA m*jp-tz-G_4* (**JP 4**), also exhibited inverse sigmoidal transition at 295 nm (Figure 5.15, and 5.16) and the T_{m} of transition was the same as that of other PNAs (all around average 40 $^{\mathrm{o}}$C). The decreasing order of T_{m} at 100 mM KCl concentration is *p7-tz-G_5* (PNA **14**) 41.0 $^{\mathrm{o}}$C < m*jp-tz-G_4* (**JP 4**) 38.9 $^{\mathrm{o}}$C < aeg-PNA-G$_5$ (PNA **17**) 37.8 $^{\mathrm{o}}$C (Table 5.5).
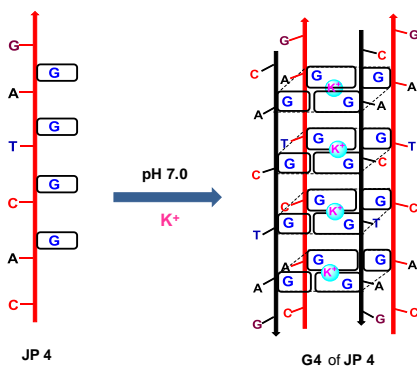
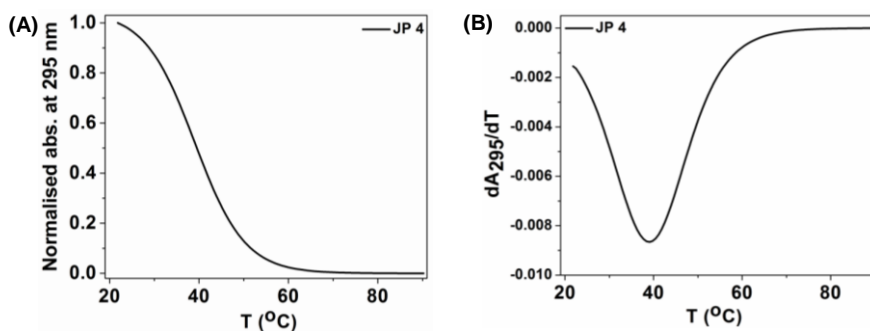**Figure 5.15** G-quadruplex of *mjp-tz-G₄* (**JP 4**)

Figure 5.16 UV-T_m of *mjp-tz-G₄* (**JP 4**) at 20 mM sodium cacodylate buffer having 100 mM KCl, pH 7.1 **(A)** normalised abs. at 295 nm **(B)** Ist derivative curve of dA_{295}/dT.

Table 5.5 UV-T_m (^oC) summary of G-quadruplex at 5 μM concentration sodium cacodylate 20 mM, pH 7.1

Entry	Peptides	295 nm KCl, 100 mM	295 nm KCl, 50 mM
1	**PNA-G₅**	37.8	38.6
2	*p7-tz-G₅*	41.0	39.8
3	*mjp-tz-G₄*	39.0	--

5.5.4 G-quadruplex studies by Variable-Temperature Circular Dichroism

For further proof of G-quadruplex formation and its stability, temperature-dependent CD experiments of *p7-tz-G₅* (PNA **14**) were done under identical conditions as that used for UV-T_m experiments. *p7-tz-G₅* (PNA **14**) in 150 μM strand concentration was prepared in 20 mM sodium cacodylate buffer containing 100 mM KCl at pH 7.1. The sample was heated to 90 ºC for 10 min, cooled slowly to room temperature over 6 h and equilibrated at 3 ºC for 24 h. The CD spectra were recorded in the temperature range 10 °C - 70 °C with heating rate of 2 °C/min. The spectra recorded between 200 nm - 320 nm are presented as average of 3 successive runs. The baseline

corresponding to buffer alone was subtracted from all spectra. The CD spectra of *p*7-*tz*-*G5* (PNA **14**) at 10 ºC present as quadruplex showed a negative band centering at 265 nm and a positive band at 288 nm and these characteristic peaks disappear at 70 $^{\circ}$C (Figure 5.17). This indicated formation of ordered structures at 10 $^{\circ}$C, the observed CD arises from quadruplex of *p*7-*tz*-*G5* (PNA **14**) and resembles that observed for antiparallel DNA quadruplexes such as d(GGTTTTGGTTTTGGTTTTGG).[20]
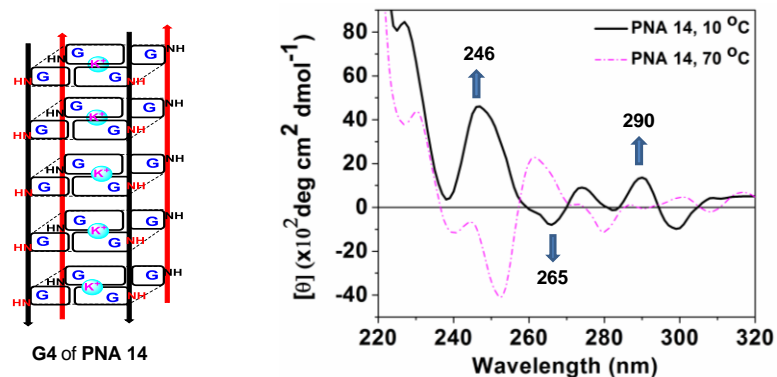

G4 of **PNA 14**

Figure 5.17 CD spectra of *p*7-*tz*-*G5* (PNA **14**, 150 μM), Sodium cacodylate 20 mM, KCl 100 mM

5.6 Self-assembly of *Janus* PNA and nanoparticles formation

Self-assembly of macromolecules generally leads into well-ordered nanostructures such as nanofibrils, nanotubes, nanospheres or vesicles, depending on the structure of the molecules and environmental conditions during assembly.[21] The relatively weak non-covalent interactions act together to form intact and well-ordered supramolecular nanostructures.[22] Gazit *et al.*[21] demonstrated the self-assembly of homo-aromatic dipeptides carrying substitutions such as halogens, nitro, phenyl, naphthyl on the phenyl ring that can modulate the noncovalent interactions. It was found that these form various morphologies such as tubular structures, nano-spherical assemblies, and fibrillar structures.

It was surmised that the self-assembling and hydrophobic properties of peptide backbone in PNA and H-bonded pairing properties of nucleobases, present on either face of *Janus* PNA may lead to interesting nanostructures (nanofibre, nanotubes, nanospheres or vesicles). *Self-complementary Janus* PNAs (*SCM*-*JP*) can in principle form sheet-like structure that may grow in 2 dimension due to complementation of amide and triazole faces (Figure 5.18) and it can be terminated at any point by addition of single face complementary *aeg* PNA/DNA/RNA, leading to possibility of regulating the self-assembly. Another characteristic feature of such small staple

like *Janus* PNAs with simultaneous presence of pi-pi inter-base and several H-bonding interactions is the synergistic complex formation from both faces. Single *Janus* PNA can complement large DNAs (such as plasmids), each face binding to distant regions, thereby bringing them together and compacting large DNA molecules. When more than one *Janus* PNA is used, it would give rise to *Janus* PNA assisted DNA origami and thereby result in programmable self-assembly to generate shape-selective nanostructures.
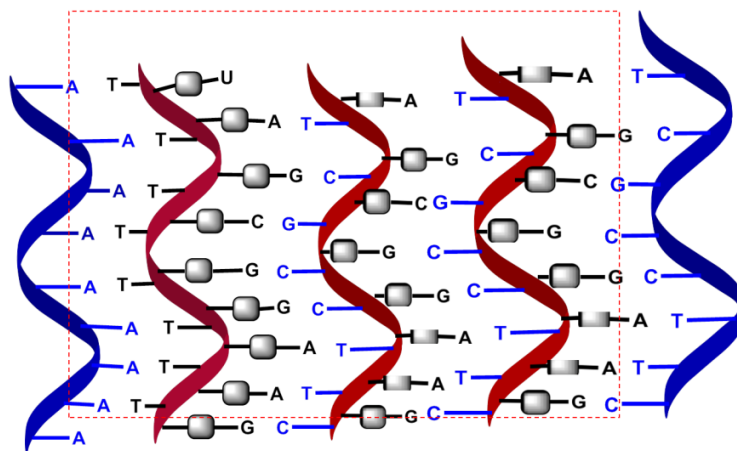

Figure 5.18 Self assembly of *self complementary Janus* PNAs

5.7. Self-assembly studies of *Janus* PNA with complementary DNA duplexes

As a first step towards that direction, this section reports on the morphology induced in DNA complexes of *homo Janus* PNAs T_7jp-tz-C_5 (**JP 1**), *chimeric Janus* PNA mjp-tz-C_5 (**JP 2**) and *self-complementary mix Janus* PNA (**JP 7**), $p7tz$-C_5 (PNA **13**), $p7tz$-G_5 (PNA **14**) and *aeg*-PNA oligomers including PNA-C_5 and PNA-G_5. PNA:DNA duplexes were constituted with appropriate complementary DNAs and all the samples were imaged at 200 μM per strand concentration in water and drop casting on silica wafer using FESEM techniques.

5.7.1 FESEM of DNA duplex and double duplex of triplex from *homo Janus* PNA T_7jp-tz-C_5 (**JP 1**)

The PNA T_7jp-tz-C_5 (**JP 1**) alone exhibited irregular sized spherical nanoparticle (500 nm) morphology with rough surface (Figure 5.19A) and when complexed with amide face complementary dA$_8$ (DNA **1**), the shape of the derived triplex C_5-tz-jpT_7:dA$_8$:T_7jp-tz-C_5 (**JP 1:DNA 1:JP 1**) changed to smoothened surface with thick outer ring surface, but inner core slightly less dense with punctation of spherical particles (Figure 5.19B). Complementation of

T_7jp-tz-C_5 (**JP 1**) from triazole face with dG_6 (DNA **2**) gave the duplex T_7jp-tz-C_5:dG_6 (**JP 1**:DNA **2**) which exhibited solid spheres with smooth surface morphology (1 μm) (Figure 5.19C). When the triplex T_7jp-tz-C_5:dA_8:T_7jp-tz-C_5 was complexed with dG_6 from two triazole faces to obtain duplex of triplex dG_6:C_5-tz-jpT_7:dA_8:T_7jp-tz-C_5:dG_6 (DNA **2**:JP1:DNA **1**:JP **1**:DNA **2**), the morphology changed to twisted nanofibers (1 μm) appearing to be spun from spherical nanoparticles (Figure 5.19D) through mass transfer. From these images, it is apparent that single duplexes and double duplexes from *homo Janus* PNA (**JP 1**) give morphological structures that are different in terms of size and shape for amide face and triazole face duplexes and for *Janus* duplex of triplex.
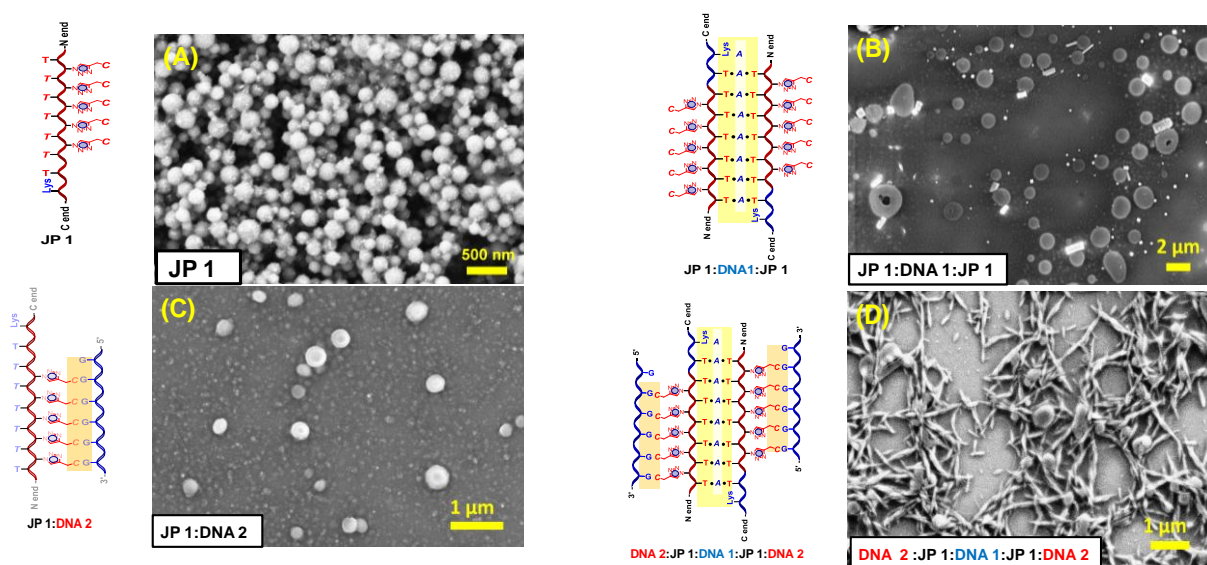

Figure 5.19 FESEM images of *homo Janus* PNA T_7jp-tz-C_5 (**JP 1**) **(A) JP 1 (B) JP 1**:DNA **1**:JP 1 (triplex) **(C) JP 1**:DNA **2** (Duplex) **(D)** dG_6:C_5-tz-jpT_7:dA_8:T_7jp-tz-C_5:dG_6 (DNA **2**:**JP 1**:DNA **1**:JP **1**:DNA **2**, double duplex of triplex).

5.7.1a Time-dependent self-assembly of Homo Janus PNA (duplex of triplex) complex: Since the nanofibers observed with the *Janus* duplex of triplex seem to be spinning out of nanospherical particles, time dependent evolution of morphology of nanofibers was studied. Immediately after drop casting, nanospherical particle (200 nm) formation was seen (Figure 5.20A). After 1 day, many of the spherical particles were observed to shrink in size and sticking to branches of nanofibers (1 μm) (Figure 5.20B), indicating that the nanofibers are indeed spun out of the initially formed spherical nanoparticles. After 10 days fully elongated but condensed fibers (10 μm) were found to be formed (Figure 5.20C).
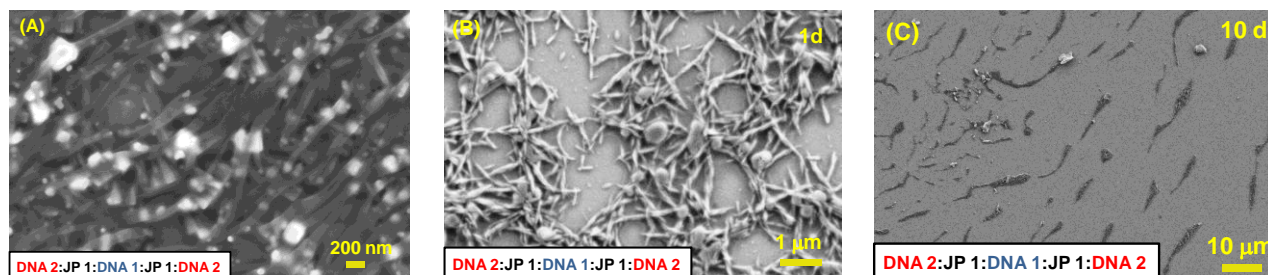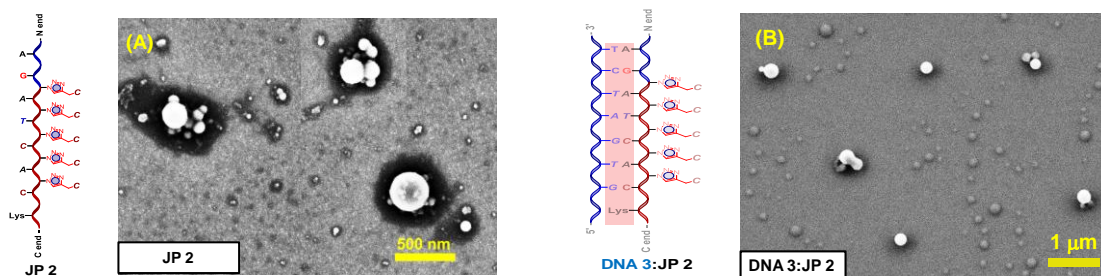
Figure 5.20 Time-dependent FESEM images of dG$_6$:*C$_5$-tz-jpT$_7$*:dA$_8$:T$_7$*jp-tz-C$_5$*:dG$_6$ (DNA **2**:**JP 1**:DNA **1**:**JP 1**:DNA **2**, double duplex of triplex) **(A)** immediate after drop cost **(B)** after 1 days **(C)** after 10 days.

5.7.2 FESEM of *chimeric Janus* PNA, JP 2 duplex with complementary DNA

The *chimeric Janus* PNA m*jp-tz-C$_5$* (**JP 2**) with mixed sequence on amide face and homo oligomeric C on triazole face shows spherical nanoparticle (500 nm) with somewhat smooth surface (Figure 5.21A) and upon forming duplex on amide face with cDNA **3** to generate the duplex DNA **3**:**JP 2**, the shape of the particle did not change much and remained smooth and spherical (1μm) (Figure 5.21B). However duplex formation of m*jp-tz-C$_5$* (**JP 2**) with dG$_6$ (DNA **2**) on triazole face, the morphology changed to elongated thick fiber (Figure 5.21C) sticking to several small particles, indicating incomplete fiber formation (1 μm). The *chimeric Janus* PNA m*jp-tz-C$_5$* (JP **2**) upon complementation on both sides to yield double duplex DNA **3**:m*jp-tz-C$_5$*:dG$_6$ (DNA **3**:**JP 2**:DNA **2**) exhibiting full fibre morphology (Figure 5.21D) finally leading to rod (500 nm) like structures. Thus the *chimeric Janus* PNA m*jp-tz-C$_5$* (**JP 2**) behaves differently in its morphology upon duplex formation with cDNA, as compared to *homo Janus* T$_7$*jp-tz-C$_5$* (JP **1**).
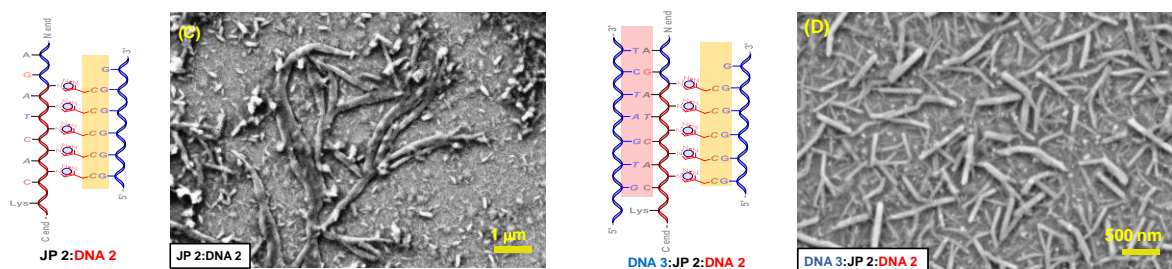
Figure 5.21 FESEM images of **(A)** *chimeric Janus* PNA m*jp-tz-C$_5$* (**JP 2**) **(B)** amide duplex with DNA **3**, DNA **3:JP 2** **(C)** triazole duplex with dG$_6$, **JP 2**:DNA **2** **(D)** Double duplex m*jp-tz-C$_5$*:dG$_6$ (DNA **3:JP 2:DNA 2**).

***5.7.2a Time-dependent assembly of duplex* m*jp-tz-C$_5$*:dG$_6$ (*JP 2):DNA 2*.** The triazole side duplex (m*jp-tz-C$_5$*:dG$_6$) (Figure 5.22A) showing the twisted tape (1 μm) morphology after 1 day was examined after 7 days and noticed that it has now become network of completely twisted long fiber bundles (1 μm) (Figure 5.22B).
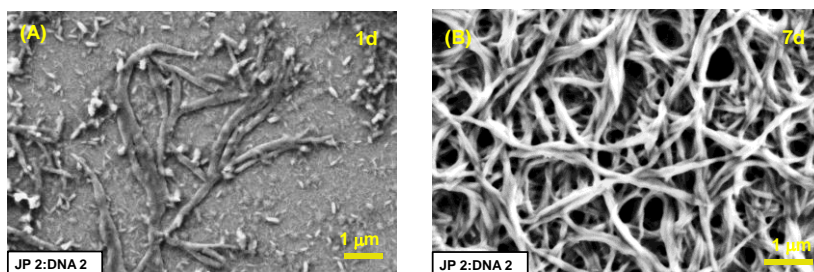

Figure 5.22 FESEM images of m*jp-tz-C$_5$*:dG$_6$ (**JP 2**:DNA **2**) in water **(A)** after 1 day and **(B)** after 7 days.

***5.7.2b Time-dependent assembly of double duplex* cDNA 3:m*jp-tz-C$_5$*:dG$_6$ (DNA 3:JP 2:DNA 2):** The self-assembly of *chimeric Janus* double duplex cDNA **3**:m*jp-tz-C$_5$*:dG$_6$ (DNA **3:JP 2:DNA 2**) showed elongated flat fiber-like morphology (500 nm) after 1 day (Figure 5.23A) and grew into long coiled noodles after 10 days (10 μm) (Figure 5.23B).
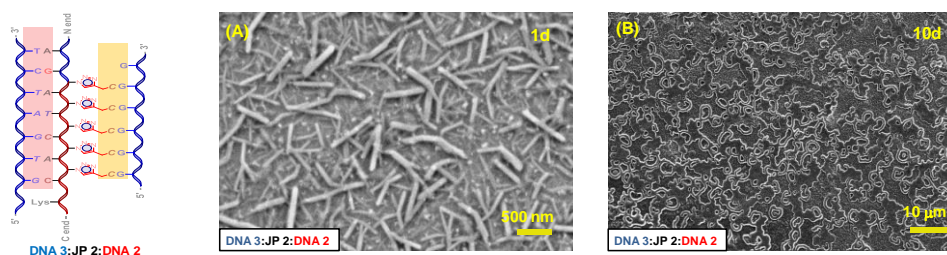

Figure 5.23 FESEM images of cDNA **3**:m*jp-tz-C$_5$*:dG$_6$ (DNA **3:JP 2:DNA 2**) in water **(A)** after 1 day **(B)** after 10 days (at 200 μM / strand concentration).

5.7.3 FESEM of *self-complementary Janus* PNA JP 7 (*SCM-JP*) with complementary DNA

The *Janus* PNA, *SCM-JP* **7** in which the amide face sequence is complementary to triazole face, should allow H bonding between the two complementary faces leading to higher order structures even without complementary DNA. Hence the *Janus* PNA, **JP 7** samples were annealed by heating up to 90 ^{o}C followed by cooling to room temperature further kept at 4 ^{o}C for 8 h. They were then drop-casted on silicon wafer, dried at room temperature for 24 h and gold coated for FESEM imaging (Figure 5.24). However, these samples showed sticky lumps (1 μm) of mass without any regular shapes. Since self assembly is enhanced with concentration, images were recorded at a higher concentration (300 mM), but no changes were seen in its morphology (1 μm) (Figure 5.24 B).
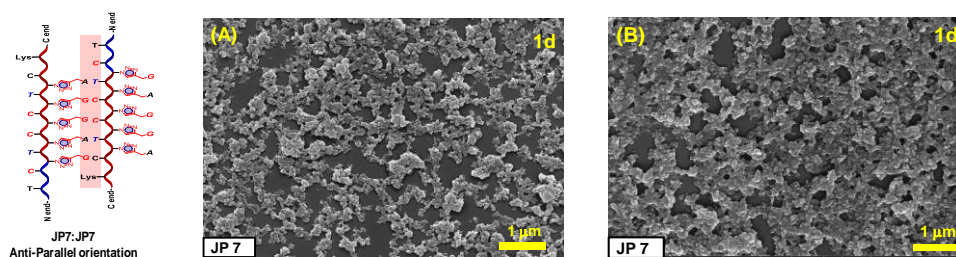

Figure 5.24 Concentration-dependent FESEM images of *SCM-JP* (**JP 7**) in water (**A**) at 200 μM and (**B**) at 300 μM.

5.7.3a Time-dependent assembly of FESEM study of double duplexes of *SCM-JP* (**JP 7**) with cDNA

The *Janus* PNA *SCM-JP* (**JP 7**) was complexed on both faces with complementary DNAs to get the double duplex DNA **8**:*SCM-JP* **7**:DNA **9**. FESEM image of this complex showed sticky clumps (1μm) (Figure 5.25A), which after 6 days showed better organized spherical nanoparticles (1μm) still growing with smaller nanoparticles sticking on the surface of bigger nanoparticles (Figure 5.25B) that slowly get intensely packed into higher density (Figure 5.25C, D).
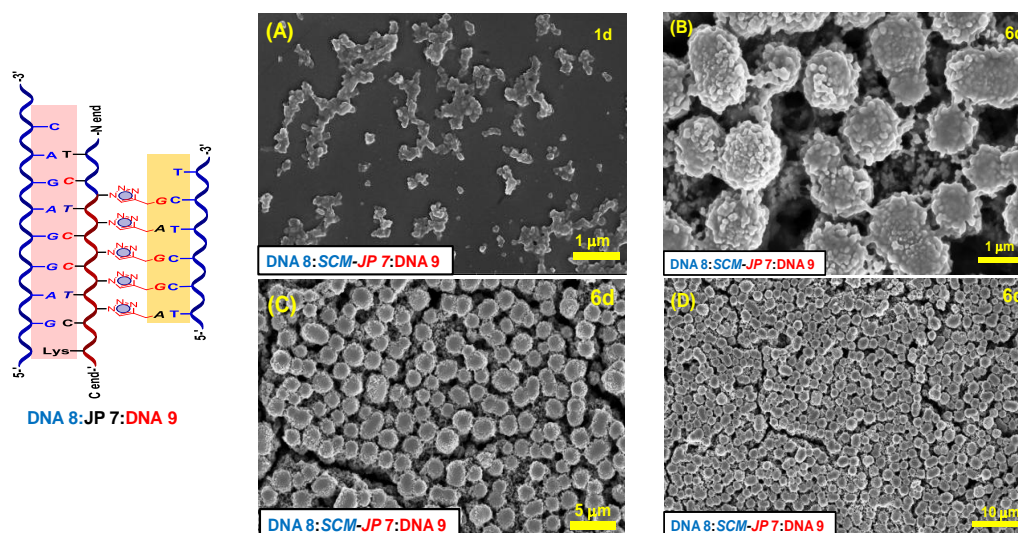
Figure 5.25 Time dependent FESEM images of *SCM*-*JP* (**JP 7**) double duplex, (DNA **8**:*SCM*-*JP 7*:DNA **9**) in water **(A)** after 1 day (DNA **8**:*SCM*-*JP 7*:DNA **9**) (**B**, **C** and **D**) after 6 days DNA **8**:*SCM*-*JP 7*:DNA **9**) at different scale bar.

5.7.4 Self-assembly studies of *aeg-tz* PNA and *aeg* PNAs and DNA duplexes

The *aeg-tz* PNAs *p7tz*-C_5 (PNA **13**), *p7tz*-G_5 (PNA **14**) and *aeg*-PNA-C_5 (PNA **16**) and *aeg*-PNA-G_5 (PNA **17**) and their complexes with appropriate cDNAs - DNA **1**, DNA **2**, DNA **3**, DNA **8**, DNA **9** and DNA **10** were examined by FESEM and image recorded at concentration of 200 μM per strands in water after 1 day incubation (Figure 5.26). dC_6 (DNA **10**) and *aeg*-PNA-G_5 showed agglomeration (200 nm) (Figure 5.26A, B), *p7tz*-G_5 (PNA **14**) (Figure 5.26C) formed amorphous rod like morphology (200 nm), PNA **17**:DNA **10** duplex was a sticky particle (1μm) (Figure 5.26D), the duplex PNA **14**:DNA **10** indicated signs of forming nanoparticles (200 nm) (Figure 5.26E). DNA **2** showed a blade like morphology (2 μM) (Figure 5.26F), PNA **16** agglomerated (200 nm) (Figure 5.26G), PNA **13** displayed round ill-defined nanoparticle (200 nm) (Figure 5.26H), the duplex DNA **2**:DNA **10** clumps of primitive nanoparticle (Figure 5.26I), duplex and PNA **16**:DNA **2** was sticky mass (2 μm) (Figure 5.26J). Overall, none of these PNAs and their duplexes with DNA exhibited any interesting morphology.

(A) DNA 10

(B) PNA 17

(C) PNA 14

(D) PNA 17: DNA 10

(E) PNA 14:DNA 10

(F) DNA 2

(G) PNA 16

(H) PNA 13
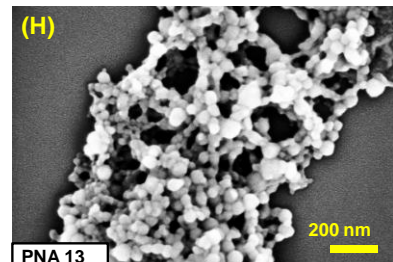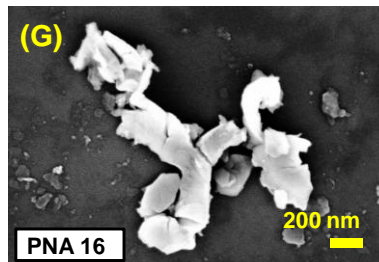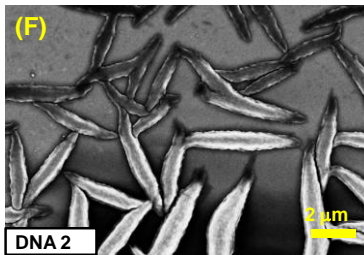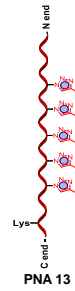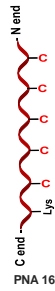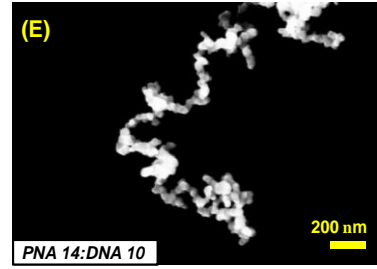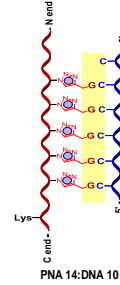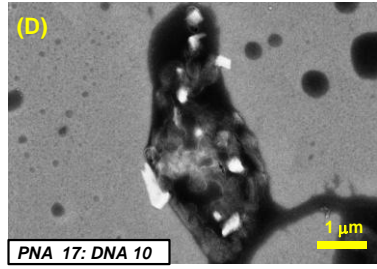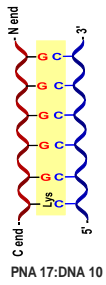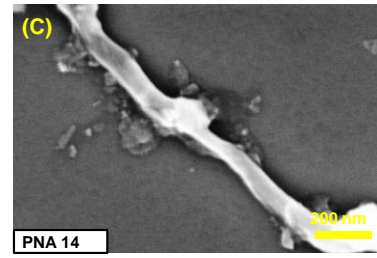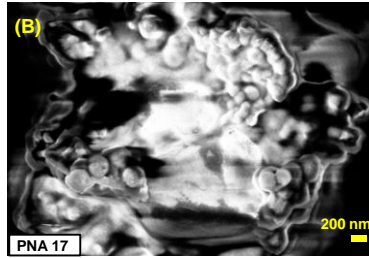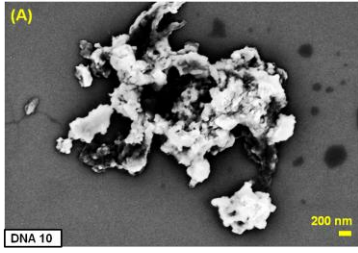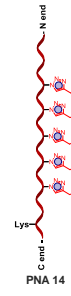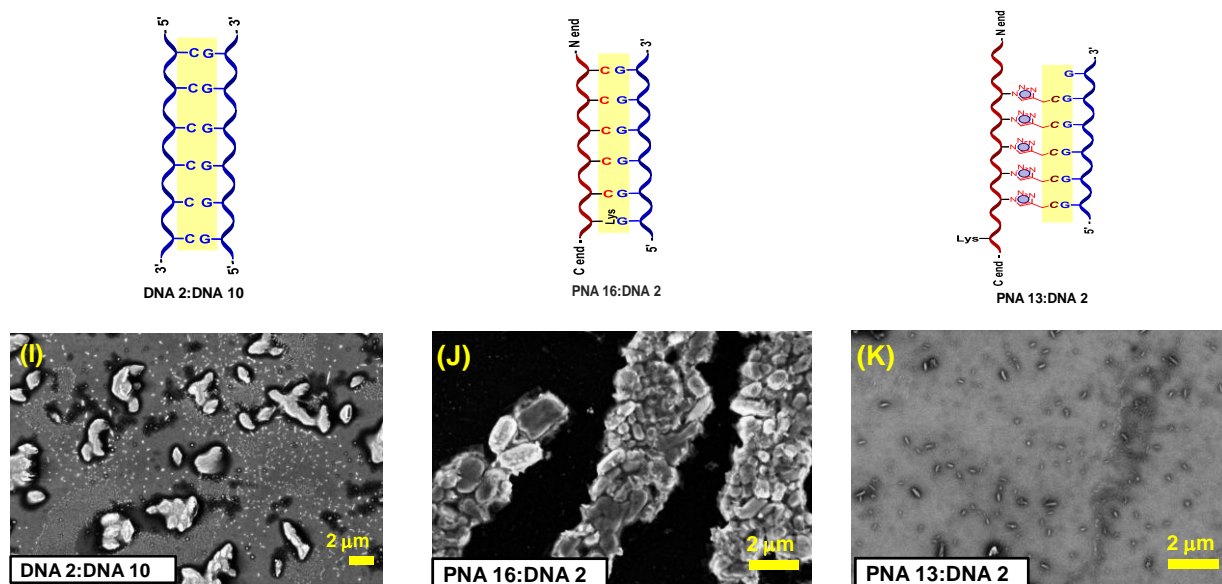
Figure 5.26 FESEM images of *aeg-tz* PNA / *aeg* PNA / DNA with complementary DNA in the water at 200 μM concentration per strand **A-K**.

5.8 Conclusion

The overall conclusions from data in this chapter involves

- pK$_a$ determination of N3 of cytidine in various PNA/DNA by pH-dependent UV-absorption shows the following order: pk$_a$ 4.03 in dC$_6$ (DNA **10**) > pK$_a$ 3.81 in *p7-tz-C$_5$* (PNA **13**) > pK$_a$ 3.60 in PNA-C$_5$ (PNA **16**).

- pH - dependent UV-T_m *i*-tetraplex studies on *Janus* PNA T$_7$*jp-tz-C$_5$* (**JP 1**) gave T_m = 57.8 ^{o}C ~ *p7-tz-C$_5$* (PNA **13**) T_m = 58.1 ^{o}C and dC$_6$ (DNA **10**) T_m = 42.6 ^{o}C till pH 5.0.

- Salt concentration-dependent UV-T_m (G$_4$ stability) at 100 mM concentration: m*jp-tz-G$_4$* (**JP 4**) T_m = 39.0 ^{o}C; *p7-tz-G$_5$* (PNA **14**) T_m = 41.0 ^{o}C and PNA-G$_5$ (PNA **17**) T_m =37.8 ^{o}C.

- CD experiment shows characteristic tetraplex signature with a band at 295 nm at 10 ^{o}C indicating formation of G$_4$ tetraplex from *p7-tz-G$_5$* (PNA **14**).

- Morphological studies of *Janus* PNA duplexes with each face complementary DNA leads to formation of various kinds of morphologies as seen with FESEM. The double duplex of *Janus* PNA with both face complementary DNA formed fiber like morphology.

  i) *Janus* PNAs T$_7$*jp-tz-C$_5$* (**JP 1**) and m*jp-tz-C$_5$* (**JP 2**) alone form spherical nanoparticles

ii) The duplex from triazole-face of $T_7jp\text{-}tz\text{-}C_5$ show solid spherically smooth surface nanoparticle (1 µM) and while that from m$jp\text{-}tz\text{-}C_5$ forms tape or network like (1 µM) morphology.

iii) The double duplex of triplex of $T_7jp\text{-}tz\text{-}C_5$ (**JP 1**) with dA$_8$ and dG$_6$ shows formation of twisted nanofibres (1 µM) and double duplex of m$jp\text{-}tz\text{-}C_5$ (**JP 2**) form fiber-like (500 nm) morphology after one day.

5.9 Summary

Janus PNAs having homo oligomeric C or G sequences on one face can form corresponding *i*- or G$_4$-tetraplexes with thermal stability similar to that from *aeg*-PNA or DNA sequences. As expected the stability of *i*-motif C$_n$ tetraplexes depend on the pK$_a$ of N3 of C and the stability of *Janus* PNA-C$_5$ is maximum at pH 5.1, close to pK$_a$ of N3 of C. Interestingly, it is found that at physiological pH, where the extent of C-protonation is less, these form duplexes, with stability that are slightly lower than that of i-tetraplexes. Similarly, *Janus* PNA-G$_4$ also forms conventional G-tetraplexes with the thermal stability that are slightly higher than that of tetraplexes from *aeg*-PNA-G$_5$. The derived supramolecular assembly studied by FESEM suggested that *Janus* PNAs form spherical nanoparticles and the double duplexes of triplexes and duplex of duplex show higher order assemblies which result in nanofiber formation. The self-complementary *Janus* PNAs however did not form well defined morphological structures from self-assembly processes. The formation of duplex of tetraplex (i-motif and G4) are interesting and have scope for further development to functionally useful architectures.
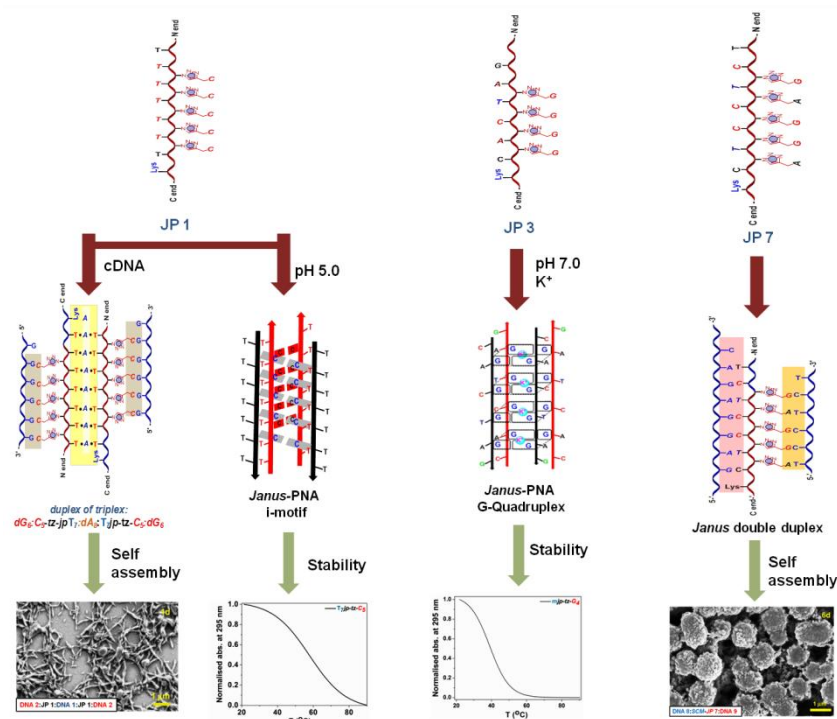
Figure 5.27 Summary of stability and self-assembly of *Janus* PNA complexes.

5.10 Experimental Procedures

5.10.1 pK$_a$ determination of oligomers by UV spectroscopy

The UV absorbance of various oligomers (DNA/PNA/amino-PNA) at 10 μM concentrations in buffers at different pHs (sodium acetate buffer 100 mM, pH 2.5 - 5.0 and sodium cacodylate 100 mM concentration, for pH 5.2 onwards used) were recorded using Shimadzu UV-2100 spectrophotometer at 25 °C. The ratios intensities at absorbance maxima of cytidine depend upon oligomer (281 – 277 nm) and protonated cytidine (273 – 269 nm) were plotted against the pH. The points obtained were fitted (Boltzmann) using Origin 8.5 software, and by Ist derivative of the fitted curve obtained the pK$_a$ value of each oligomer.

5.10.2 UV-*T*$_m$

Thermal denaturation studies of duplexes and triplexes were performed on Cary 300 Bio UV-Visible Spectrophotometer. The samples were heated at 95 °C for 10 min and cooled slowly to ambient temperature. The solution was moved to cold room and kept in freeze at (3 °C) for 24h. The variation of absorbance at 295 nm for i-motif and G-quadruplex study were measured

as a function of temperature from 20 ^{o}C - 90 ^{o}C, with absorbance recordings at every 1 ^{o}C increment. Each melting experiment was repeated at least twice and T_m values were determined from first derivative curves of normalized UV-temp plots. The concentrations of oligonucleotides were calculated on the basis of absorbance at 260 nm, using molar extinction coefficients of the corresponding nucleobases: C = 7.3 cm^2/μmol; G = 11.7 cm^2/μmol.[23]

5.10.3 Circular Dichroism (CD) spectroscopy

For CD study, the samples were prepared in sodium cacodylate buffer containing 100 mM KCl, at pH 7.2. *p*7-*tz*-*G$_5$* (PNA **14,** 150 μM) was heated to 90 ℃ for 10 min, slowly cooled to room temperature over 6 h and equilibrated at 3 ℃ for 24 h. Spectra were recorded from 10 °C to 70 °C at every 10 ^{o}C intervals, scanned from 320 nm to 200 nm and presented as average of 3 successive runs. Baseline corresponding to buffer alone were subtracted from all recorded spectra.

5.10.4 Field Emission Scanning Electron Microscopy (FESEM)

FESEM images were recorded using a Zeiss Ultra Plus scanning electron microscope and the samples were prepared by drop on freshly cut silicon wafers and coated with gold and samples were made in mili-Q water to avoid the effect of inorganic salts from the background. The elemental composition of the peptide as well as DNA was confirmed form EDAX data obtained from the SEM images (SI). For each experiment 200 μM per strands concentration of oligomers in water used for FESEM.

5.11 References

1. Watson, J. D.; Crick, F. H. C. *Nature* **1953**, *171*, 737-738.

2. (a) Bacolla, A.; Wells, R. D. *Mol Carcinog.* **2009**, *48*, 273-285. (b) Wang, G.; Vasquez, K. M. *Mutat. Res.* **2006**, *598*, 103-119.

3. Dhakal, S.; Yu, Z.; Konik, Cui, R. Y.; Koirala, D.; Mao, H. *Biophys J.* **2012**, *102*, 2575-2584.

4. Laisne, A.; Pompon, D.; Leroy, J.-L. *Nucleic Acids Res.* **2010**, *38*, 3817-3826.

5 Zeraati, M.; Langley, D. B.; Schofield, P.; Moye, A. L.; Hughes, R. W. E.; Bryan, T. M.; Dinger, M. E.; Christ, D. *Nat. Chem.* **2018**, *10*, 631-637.

6. a) Gueron, M.; Leroy, J.-L. *Curr. Opin. Struct. Biol.* **2000**, *10*, 326-331; b) Choi, J.; Majima, T. *Chem. Soc. Rev.* **2011**, *40*, 5893-5909.

7. Gehring, K.; Leroy, J.-L.; Gueron, M. *Nature* **1993**, *363*, 561-565.

8. Phan, A. T.; Kuryavyi, V.; Patel, D. J. *Curr. Opin. Struct. Biol* **2006**, *16*, 288-298.

9. (a) Guo, K.; Pourpak, A.; Beetz-Rogers, K.; Gokhale, V.; Sun, D.; Hurley, L. H. *J. Am. Chem. Soc.* **2007**, *129*, 10220-10228. (b) Leroy, J.-L.; Gueron, M.; Mergny, J.-L.; Helene, C. *Nucleic Acids Res.* **1994**, *22*, 1600-1606. (c) Xu, Y.; Sugiyama, H. *Nucleic Acids Res.* **2006**, *34*, 949-954; d) Dhakal, S.; Schonhoft, J. D.; Koirala, D.; Yu, Z.; Basu, S.; Mao, H. *J. Am. Chem. Soc.* **2010**, *132*, 8991-8997.

10. Kendrick, S.; Hurley, L. H. *Pure Appl. Chem.* **2010**, *82*, 1609-1621.

11. Dhakal, S.; Lafontaine, J. L.; Yu, Z.; Koirala, D.; Mao, H. *PLoS ONE* **2012,** *7*, e39271.

12. Sharma, N. K.; Ganesh K. N. *Chem. Commun.* **2005**, 4330-4332.

13. Xue, Y.; Kan, Z.-y.; Wang, Q.; Yao, Y.; Liu, J.; Hao, Y.-h.; Tan, Z. *J. Am. Chem. Soc.* **2007**, *129*, 11185-11191.

14. Gellert, M.; Lipsett, M. N.; Davies, D. R. *Proc. Natl. Acad. Sci. U. S. A.* **1962**, *48*, 2013-2018.

15. Pinnavaia, T. J.; Marshall, C. L.; Mettler, C. M.; Fisk, C. L.; Miles, H. T.; Becker, E. D. *J. Am. Chem. Soc.* **1978**, *100*, 3625-3627.

16. *Fundamentals of Quadruplex Structres* (Eds.: Neidle, S.; Balasubramanian S., *RSC,* **2006**, pp. 1-30.

17. (a) Smargiasso, N.; Rosu, F.; Hsia, W.; Colson, P.; Baker, E. S.; Bowers, M. T.; De Pauw, Gabelica, E. V. *J. Am. Chem. Soc.* **2008**, *130*, 10208-10216. (b) Huppert, J. L. *The FEBS Journal* **2010**, *277*, 3452-3458. (c) Risitano, A.; Fox, K. R. *Nucleic Acids Res.* **2004**, *32*, 2598-2606. (d) Hazel, P.; Huppert, J.; Balasubramanian, S.; Neidle, S. *J. Am. Chem. Soc.* **2004**, *126*, 16405-16415. (e) Miller, M. C.; Buscaglia, R.; Chaires, J. B.; Lane, A. N.; Trent, J. O. *J. Am. Chem. Soc.* **2010**, *132*, 17105-17107. (f) B. Heddi, A. T. Phan, *J. Am. Chem. Soc.* **2011**, *133*, 9824-9833. (g) Hansel, R.; Lohr, F.; Foldynova-Trantirkova, S.; Bamberg, E.; Trantirek, L.; Dotsch, V. *Nucleic Acids Res.* **2011**, *39*, 5768-5775. (h) Guedin, A.; Gros, J.; Alberti, P.; Mergny, J.-L. *Nucleic Acids Res.* **2010**, *38*, 7858-7868.

18. Vernille, J. P.; Kovell, L. C.; Schneider, J. W. *Bioconj. Chem.* **2004**, *15*, 1314-1321.

19. Sharma, N. K.; Ganesh, K. N. *Org. Biomol. Chem.* **2011**, *9*, 725-729.

20. (a) Krishnan-Ghosh, Y.; Stephens, E.; Balasubramanian, S. *J. Am. Chem. Soc.* **2004**, *126*, 5944-5945. (b) Dapic, V.; Abdomerović, V.; Marrington, R.; Peberdy, J.; Rodger, A.; Trent, J. O.; Bates, P. J. *Nucleic Acids Res.* **2003**, *31*, 2097-2107.

21. (a). Zhang, S. *Nat. Biotechnol.* **2003**, *21*, 1171-1178. (b). Song, Y.; Challa, S.R.; Medforth, *et al. Chem.Commun.* **2004**, *9*, 1044-1045. (c). Reches, M.; Gazit, E. *Nat. Nanotechnol.* **2006**, *1*, 195-200. (d). Reches, M.; Gazit, E. *Phys. Biol.* **2006**, *3*, S10-S19

22. Reches, M.; Gazit, E. *Curr. Nanosci.* **2006**, *2*, 105-111.

23. D. R. Jain, L. Anandi V, M. Lahiri, Ganesh, K. N. *J. Org. Chem.* **2014**, *79*, 9567-9577.

Summary of thesis and future outlook

This thesis presents a new and novel type of PNA structure termed "*Janus* PNA" in which each PNA unit consists of an additional side chain at C_α position which carries a nucleobase through a triazole linker. Thus the PNA backbone of *Janus* PNA is endowed with two nucleobase sequences on two sides of the backbone termed "amide face" and "triazole face". Therefore, the *Janus* PNA can hybridize with two cDNAs, one from the amide face and the other from triazole face leading to double duplex formation. Based on the nature of sequences on each face they have been classified into different types: *homo Janus* PNAs (T_7jp-tz-C_5, homo oligomeric sequences on both face), *chimeric Janus* PNA (mjp-tz-C_5, mixed sequences on amide face and homo oligomeric sequence on triazole face), *hetero Janus* PNA (mjp-tz_5-*CCACG*, mixed sequence on both face) and *self-Complementary Mix Janus* PNA (*SCM*-*JP*). These were synthesized on solid phase manually using appropriate protected monomers, cleaved from resin, purified and characterized by mass spectra.

The complex formation of different Janus PNAs was studied by UV-T plots and observation of single sigmoidal transitions suggested successful duplex/triplex formation. The triazole PNA sequences having only nucleobases conjugated via C_α-side chain via triazole linker and devoid of tertiary amide linked triazole side chain (as in standard *aeg*-PNA) formed duplex with cDNA, indicating their individual capability to base pair with cDNA. The different *Janus* PNAs having base sequences on both amide and triazole faces were hybridized with appropriate cDNAs to generate duplexes (amide face/triazole face), triplexes and double duplex of triplex (with *homo Janus* PNAs), and double duplexes (*chimeric* and *hetero Janus* PNAs). The T_ms of each of these duplexes were measured by temperature dependent UV absorbance. In isolated individual duplexes with one cDNA binding, it was found that the triazole side duplex always had a higher T_m than amide face duplex. In complexes composed in presence of both cDNAs, double duplexes were observed with simultaneous binding on both sides to duplexes on both faces. These gave two transitions and hence two T_ms corresponding to melting of each duplex. Two interesting observations: (i) The T_m of both amide and triazole duplex in the double duplex were much higher than analogous T_ms in in individual isolated duplexes and (ii) the triazole duplexes always had higher T_m than amide face duplex. To establish the sequence dependence of base pair formation, all individual and Janus double duplexes were constituted from mismatched

DNA sequences. These exhibited lower T_ms than the perfect duplexes indicating the sequence dependence of duplex formation.

The conformational aspects of duplex and triplex formation from *Janus* PNA was studied by CD spectroscopy. The derived duplexes and triplexes generated by stochiometric additions of *Janus* PNA and DNA components exhibited CD profiles characteristic of PNA:DNA duplexes and triplexes confirming similar conformation. The double duplex of Janus PNA can be generated in two ways depending on the order of duplex formation. CD studies of sequential formation of duplexes indicated that irrespective of the order, the final duplex always has same conformation.

The thermodynamic parameters for binding reactions of different *Janus* PNAs with cDNA were obtained from Isothermal Titration Calorimetry. The stoichiometric and co-operative nature of binding reactions were indicated by sigmoidal nature of binding isotherm. The enthalpy and free energy of various binding reaction were negative and quantitatively similar to that seen in standard *aeg*-PNA: DNA complexes. The stoichiometry of complexation in different experiments computed in terms of ratio of number of DNA and PNA bases involved in base pair formation and corresponding to the proposed duplex / triplex / double duplexes matched with that obtained by the experiment assuming one-site / two site binding models. This further supported the formation of different hybrids through specific base pairing compositions to yield various complexes. The combined biophysical results established the proof of concept of design of *Janus* PNAs and demonstrated their ability to simultaneously bind to cDNAs from both sides to form double duplexes.

The formation of higher ordered structures such as *i*-motif from *Janus* PNA (T$_7$*jp-tz-C$_5$*) and tetraplexes from *Janus* PNA (m*jp-tz-G$_4$*) were also examined under relevant pH and salt conditions. These do indeed show formation of i-motif and G$_4$-tetraplex structures with stability as much as the parent PNA. The self-assembly of various *Janus* PNAs and their complexes with cDNAs were investigated through their morphological structures by FESEM. Very distinct morphological features ranging from spherical particles to fibers were noticed in each case clearly indicating specific complex formations through base pairing.

Nature of base pairing: The overall work in the thesis clearly confirms interaction of *Janus* PNAs from both sides to form sequence specific base pairing through complementary H-bonding. Based on all literature of H-bonding among natural nucleobases, it is tempting to suggest that all base pairing observed in this work are Watson-Crick base pairing. The other

not so prevalent form is Hoogsteen base pairing. There are not many cases where isolated Hoogsteen base pairing occurs in solution though it is well characterized in crystals. They are much weaker than canonical Watson-Crick base pairing and in DNA and PNA triplexes, they occur only when the bases are paired by Watson-Crick H-bonding on one face and not in isolation. These facts clearly point out that in *Janus* PNA, all base pairing seen among various duplexes are Watson-Crick in motif. It can only be established by NMR spectroscopic studies or X-ray crystal structures, which need enough materials. Hopefully in future this can be achieved. Hence in this thesis it has been refrained from the use of "Watson-Crick H-bonding."

Future outlook: If base pairing in *Janus* PNA is by Watson-Crick H-bonding, which is most possibly true, the double duplex formation by *Janus* PNAs by simultaneous recognition of a single strand *Janus* PNA by two complementary DNA strands, both by sequence specific Watson-Crick H-bonding is the most novel aspect of the work. This would certainly open up new biological applications for *Janus* PNAs.

(i)   When *Janus* PNAs form i-motif or G_4-tetraplexes, on the backbone another set of sequences are available for further growth of self-assembly for higher order structures as shown in the following figures. One can generate tetraplex of tetraplexes and contiguous i-motif structures. If the sequence is a mixed sequence, one can generate duplexes of i-motif or tetraplexes. Given the high T_m and robustness of these structures, these may have interesting material properties.
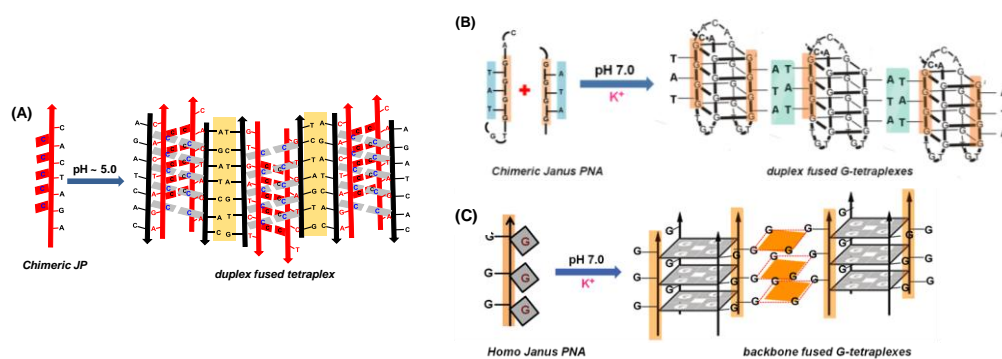

Figure 1: *Janus* PNA i-motif/G_4-quadruplex

(ii)   *Janus* PNAs can also replace DNA staples in DNA origami. Engineering ordered staples to complement specific regions in plasmid DNA, they can systematically generate various shaped structures by DNA origami.
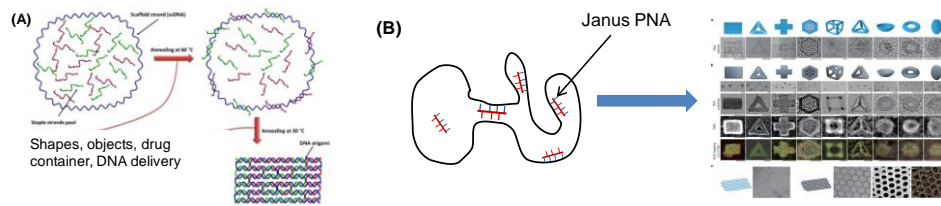
Figure 2: *Janus* PNA assisted DNA origami

(iii)   There could be specific biological applications where a single probe may be necessary to simultaneously target two different genes and *Janus* PNAs would be very versatile. Availability of such molecules may generate new applications as well. One potential application would be in *Janus* PNA or bundles or self-aggregates simplifying the enzyme machinery in eukaryotic transcription, with reference to enhancer elements. *Janus* PNAs can "in principle" bring together the relevant parts of distantly located DNA regions together to modulate gene expression. This could be a far-fetched application, availability of *Janus* PNA type of molecules may generate simpler applications.
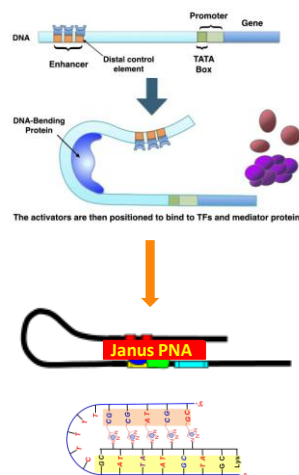

Figure 3: *Janus* PNA for eukaryotic transcription regulation

(iv)   Recently Danith Ly *et. al.* have published a nice application of control of disease related RNA hairpins using Janus bases. *Janus* PNAs could also be employed in such application.